

# **SAE 3.03 Description et prévision de données temporelles – 2025/2026**

Rapport d'activité :  
Indice de chiffre d'affaires- Location de  
logements (Janv. 2010-Déc. 2015)

## Sommaire

Introduction .....	3
Partie 1 : Présentation des données.....	4
Source des données : .....	4
Indice de chiffre d'affaires - Location de logements (NAF rév. 2, niv. sous-classe poste 68.20A) (INSEE) .....	4
Choix du modèle de décomposition .....	5
Partie 2 : Décomposition, Modélisation et Validation .....	6
Décomposition de la série .....	6
Coefficients Saisonniers et Série CVS .....	7
Modélisation Linéaire de la Tendence .....	9
Validation du Modèle par l'Analyse des Résidus.....	10
Partie 3 : Prévisions et Comparaison .....	12
Prévisions à Court Terme par Modélisation Linéaire .....	12
Prévisions par Lissage Exponentiel (Holt-Winters) .....	13
Comparaison des Performances et Modèle Retenu .....	14
Prévisions Finales .....	14
Conclusion : .....	15



# Introduction

**Contexte** : L'importance du secteur de la location de logements.

**Problématique** : Comment a évolué ce secteur entre 2010 et 2015 ? Présente-t-il des schémas saisonniers ? Est-il possible de prévoir son évolution à court terme ?

Ce rapport s'appuie sur le jeu de données de l'**Indice de Chiffre d'Affaires (ICA) pour la location de logements**, couvrant la période de **janvier 2010 à décembre 2015**, afin de modéliser et de prédire son évolution temporelle.

**Objectif du rapport** : Répondre aux exigences de la SAE, c'est-à-dire décomposer la série, la modéliser et effectuer des prévisions.

**Annonce du plan** :

Partie 1 : Présentation et Analyse Descriptive des Données

Partie 2 : Décomposition, Modélisation et Validation

Partie 3 : Prévisions et Comparaison



# Partie 1 : Présentation des données

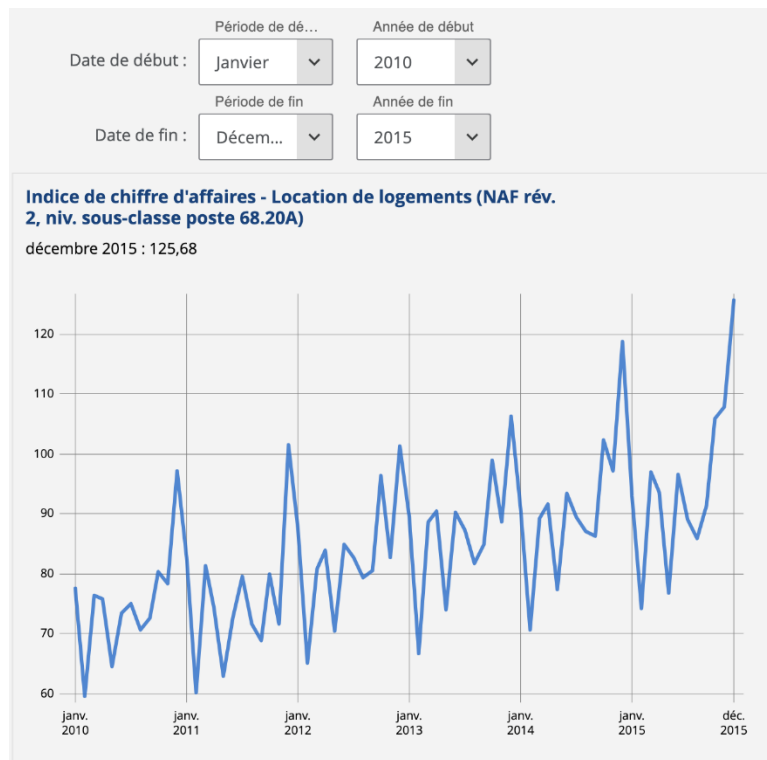
Source des données :

[Indice de chiffre d'affaires - Location de logements \(NAF rév. 2, niv. sous-classe poste 68.20A\) \(INSEE\)](#)

Champ: France

Sélection de la période : Janvier 2010 – Décembre 2015 (6 périodes)

Fréquence : Mensuelle



Axe Y : Indice.

Axe X : Temps (de Janv. 2010 à Déc 2016).

Première valeur : 77.65 (Janv. 2010).

Dernière valeur : 125.68 (Déc. 2015).



## **Premières observations :**

En regardant le graphique et en vérifiant les données brutes, l'Indice présente trois caractéristiques majeures :

1. **Tendance croissante** : La valeur de l'indice est en hausse nette entre 2010 (indice min. 59,53) et 2015 (indice max 125,68). Le secteur est en croissance sur cette période.
2. **Saisonnalité très nette** : La série présente un cycle annuel stable et fort :
  - Les creux sont systématiquement atteints en début d'année (janvier ou février).
  - Le pic annuel absolu est toujours atteint en décembre, après une forte activité estivale (juillet/août). Cette forte performance en fin d'année est un trait caractéristique de cette série.
3. **Modèle multiplicatif** : L'écart absolu entre le pic annuel et le creux d'hiver augmente en même temps que la tendance générale monte. Cela signifie que le choix d'un modèle multiplicatif est le plus pertinent, car les variations saisonnières sont proportionnelles au niveau de l'indice.

## Choix du modèle de décomposition

Nous allons procéder à la décomposition de la série en ses trois composantes : la Tendance, la Saisonnalité et le Résidu.

Le choix du modèle se fait entre deux options principales :

- **Modèle Additif** : Les composantes s'ajoutent. L'amplitude des variations saisonnières est constante dans le temps.

$$Y_t = T_t + S_t + \epsilon_t$$



- **Modèle Multiplicatif** : Les composantes se multiplient. L'amplitude des variations saisonnières est proportionnelle au niveau de la tendance.

$$Y_t = T_t \times S_t \times \epsilon_t$$

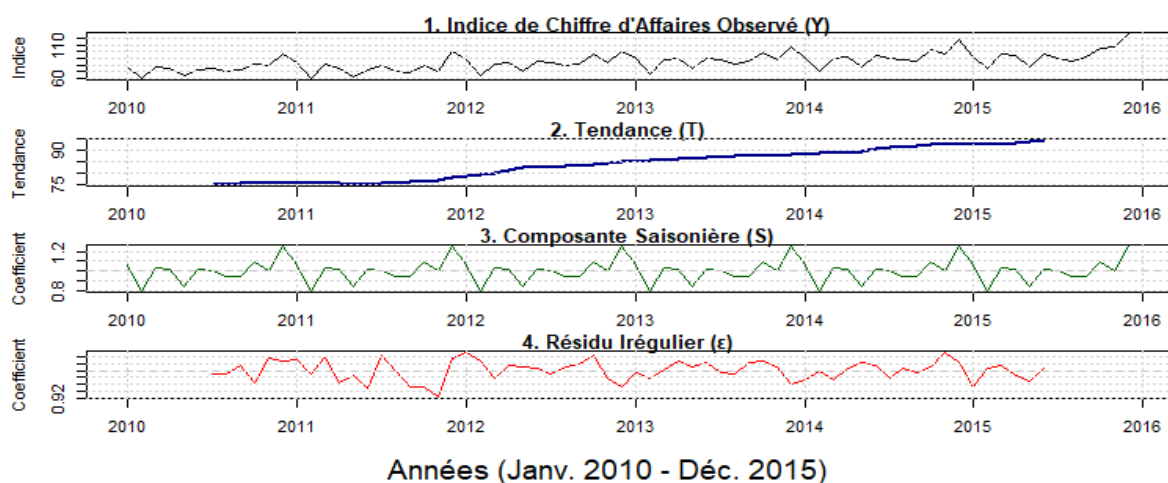
Comme nous l'avons constaté lors de l'analyse visuelle (Page 4), l'amplitude des variations saisonnières augmente en même temps que la tendance générale croît. C'est pourquoi nous choisissons de travailler avec le **Modèle Multiplicatif** pour la suite de l'analyse.

## Partie 2 : Décomposition, Modélisation et Validation

### Décomposition de la série

Nous appliquons la décomposition multiplicative choisie.

#### Décomposition Multiplicative de l'Indice de Chiffre d'Affaires (ICA)



(Utilisation de l'intelligence artificielle pour le code et le design de ce graphique pour bien représenter l'idée)



**Analyse de la Tendance ( $T$ ) :**

Elle confirme une **croissance quasi-linéaire** de l'activité de fond.

**Analyse de la Saisonnalité ( $S$ ):**

Il montre un cycle très stable, avec des valeurs **supérieures à 1** en fin d'année (Décembre) et en été, et des valeurs **inférieures à 1** en début d'année (Janvier/Février).

**Analyse des Résidus ( $\epsilon$ ):**

Les valeurs sont centrées autour de 1, mais on observe une **volatilité notable** (fluctuations irrégulières), en particulier vers la fin de la série (2015). Cela indique que le modèle de décomposition simple **ne capture pas tout** l'aspect irrégulier, ce qui sera crucial pour la validation (Page 9).

- Cette décomposition permet de calculer les coefficients saisonniers (Page 7) et de modéliser la tendance (Page 8).

## Coefficients Saisonniers et Série CVS

Le modèle multiplicatif nous a permis d'estimer les coefficients saisonniers ( $S$ ), qui mesurent l'influence de chaque mois sur l'indice moyen. Un coefficient supérieur à 1 indique une activité supérieure à la moyenne annuelle, et inversement pour un coefficient inférieur à 1.

## 1. Quantification de la Saisonnalité

Mois	Coefficient ( $S_t$ )	Interprétation
Janvier	1.0554355	Activité <b>5,54 % au-dessus</b> de la moyenne.
Février	0,7967	<b>Creux :</b> Activité <b>20,33 % en dessous</b>



		de la moyenne.
Mars	1,0311	
Avril	1,0185	
Mai	0,8446	
Juin	1,0152	
Juillet	1,0035	
Août	0,9426	
Septembre	0,9458	
Octobre	1,0958	
Novembre	0,9985	
Décembre	1,2523	<b>Pic Annuel : Activité 25,23 % au-dessus de la moyenne.</b>

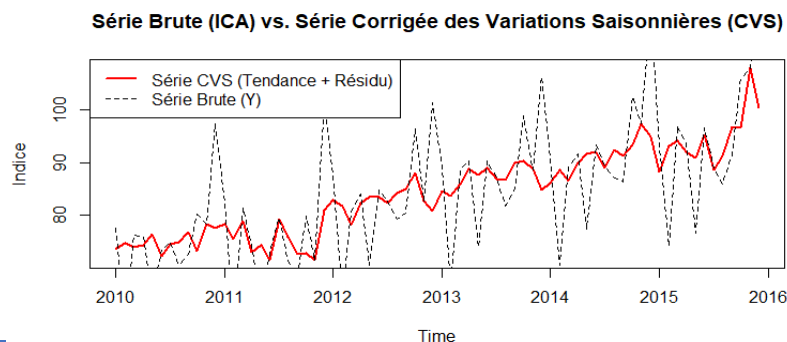
*Note sur janvier* : Contrairement à l'analyse visuelle (creux), Janvier apparaît ici comme un mois **au-dessus de la moyenne**, tandis que **février** est le véritable mois le plus bas. Ceci est la valeur moyenne *globale* de l'indice des 6 années. Le pic de **décembre** est de loin la période la plus forte.

## 2. Série Corrigée des Variations Saisonnières (CVS)

La Série Corrigée des Variations Saisonnières ( $y_t^{\{CVS\}}$ ) est obtenue en **retirant l'effet saisonnier** ( $\hat{S}_t$ ) de la série brute ( $y_t$ ).

Dans le cadre de notre modèle multiplicatif, la formule appliquée est :

$$y_t^{\{CVS\}} = \frac{(y_t)}{(1 + \hat{S}_t)}$$



### Commentaire sur la CVS :

La superposition montre que la série CVS est beaucoup plus lisse que la série brute. Toutes les fluctuations récurrentes ont disparu, ne laissant apparaître que l'**évolution de fond (la tendance)** et les chocs imprévus. La série CVS confirme la **nette croissance** de l'activité du secteur, que nous allons maintenant chercher à modéliser.

### Modélisation Linéaire de la Tendance

Nous avons observé sur la série  $y_t^{\{CVS\}}$  une croissance qui semble constante. Nous choisissons donc de modéliser la tendance par une **régression linéaire simple**, avec la variable temps  $t$  comme unique prédicteur :

$$\hat{T}_t = \beta_0 + \beta_1 \times t$$

L'exécution de la régression linéaire sur la série  $y_t^{\{CVS\}}$  donne les résultats suivants :

$$R^2 \text{ ajusté} : 0,8500$$

$$\text{Ecart-type résiduel} : 3,124$$

$\beta_0 = 71,40$  : C'est la valeur estimée de la tendance au début de la série (Janvier 2010, théoriquement à  $t = 0$ ).

$\beta_1 = 0,356$  : Il indique que l'Indice désaisonnalisé augmente en moyenne de **0,356 point par mois**. Ceci confirme une croissance nette et stable du marché de la location de logements sur la période étudiée.

### Analyse et Interprétation du Modèle

#### Équation Modélisée

L'équation de la tendance linéaire ajustée est :

$$\hat{T}_t = 71,40 + 0,356 \times t$$

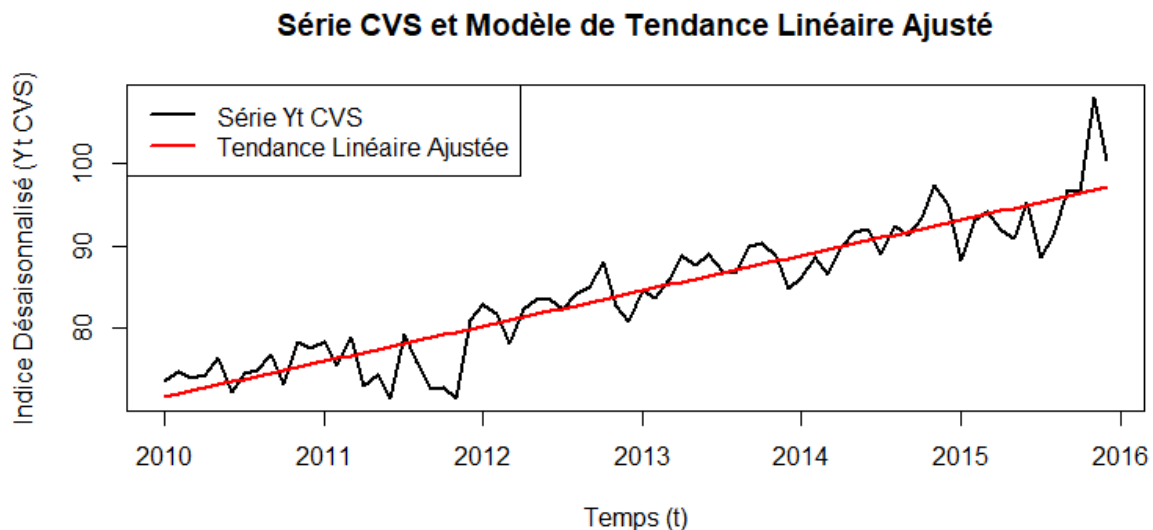


## Qualité et Validation du Modèle Linéaire

- **Coefficient de Détermination ( $R^2$  ajusté) :** Avec un  $R^2$  ajusté de **0,8500 (soit 85%)**, le modèle linéaire simple (la droite) explique une part très élevée de la variance de la série désaisonnalisée  $y_t^{\{CVS\}}$

**Conclusion :** Le modèle linéaire est une excellente représentation de la tendance de fond.

Tendance Estimée vs. Tendance Modélisée

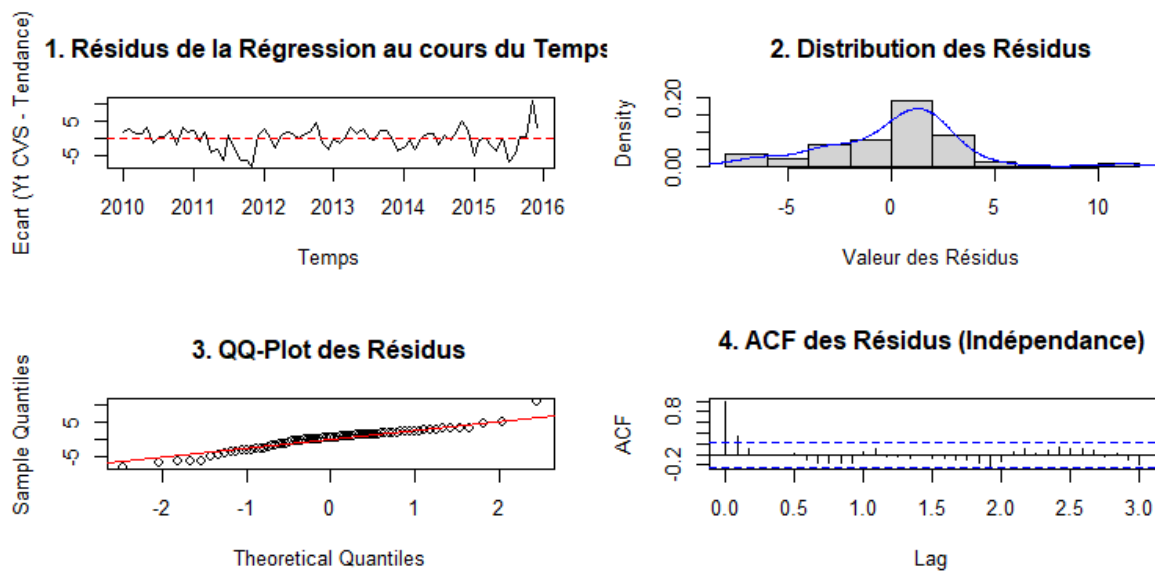


**Commentaire :** Le graphique montre clairement que la droite de régression linéaire (en rouge) s'ajuste très bien à la série  $y_t^{\{CVS\}}$ . Cela valide notre choix d'utiliser cette équation simple pour les prévisions de la tendance.

## Validation du Modèle par l'Analyse des Résidus

La validité du modèle de décomposition repose sur l'hypothèse que le terme résiduel doit être un bruit aléatoire, non corrélé et normalement distribué. Nous analysons ici les résidus du modèle de tendance linéaire.





### Observations et Validation

- **Résidus au cours du Temps (1) :**
  - Les résidus sont centrés autour de zéro, ce qui est correct.
  - On n'observe pas de tendance ni de saisonnalité restante. Cependant, la **volatilité semble augmenter** vers la fin de la période (2014-2015), suggérant que le modèle est moins précis sur les dernières années.
- **Normalité (2 & 3) :**
  - L'histogramme et le QQ-Plot montrent une distribution des résidus qui est **globalement normale**. Les points sur le QQ-Plot suivent bien la droite de référence, sauf aux extrémités.
- **Auto-corrélation (4) :**
  - La majorité des barres se trouvent dans la zone bleue de confiance

### Repérage des Valeurs Mal Ajustées (

L'analyse des résidus bruts (1) et la sortie du R montrent les valeurs mal ajustées:

- **Résidu Minimum :** -7,9633
- **Résidu Maximum :** 11,2581



La valeur maximale de **11,26** correspond très probablement à la fin de l'année **2015**, où la série  $y_t^{\{CVS\}}$  atteint un pic très élevé. Cela signifie que le modèle linéaire a **fortement sous-estimé** la croissance de l'ICA à la fin de la période, et que cet événement est considéré comme une forte irrégularité non expliquée par la tendance ou la saisonnalité.

## Partie 3 : Prévisions et Comparaison

### Prévisions à Court Terme par Modélisation Linéaire

L'approche paramétrique combine la prévision de la tendance linéaire et la réapplication des coefficients saisonniers :

$$\hat{Y}_{t+h} = \hat{T}_{t+h} \times \hat{S}_{\text{saison}(t+h)}$$

Le modèle prévoit une poursuite de la croissance (pente de 0,356 point par mois) en respectant les pics habituels de fin d'année.

Prévisions (Modèle Paramétrique) pour 2016 :

Mois	Tendance $\hat{T}_t$	Coeff. $\hat{S}_t$	Prévision $\hat{Y}_t$
<b>Janv. 2016</b>	97,38	1,0554	<b>102,78</b>
<b>Fév. 2016</b>	97,74	0,7967	<b>77,86</b>
<b>Mars 2016</b>	98,09	1,0311	<b>101,15</b>
<b>Avril 2016</b>	98,45	1,0185	<b>100,27</b>



Mois	Tendance $\hat{T}_t$	Coeff. $\hat{S}_t$	Prévision $\hat{Y}_t$
<b>Mai 2016</b>	98,80	0,8446	<b>83,45</b>
<b>Juin 2016</b>	99,16	1,0152	<b>100,66</b>
<b>Juil. 2016</b>	99,52	1,0035	<b>99,86</b>
<b>Août 2016</b>	99,87	0,9426	<b>94,14</b>
<b>Sept. 2016</b>	100,23	0,9458	<b>94,80</b>
<b>Oct. 2016</b>	100,58	1,0958	<b>110,21</b>
<b>Nov. 2016</b>	100,94	0,9985	<b>100,79</b>
<b>Déc. 2016</b>	101,29	1,2523	<b>126,85</b>

Le modèle paramétrique anticipe une **poursuite de la croissance** de la tendance au-delà de 100, ainsi qu'un respect du cycle saisonnier :

- Le **creux** de l'activité est toujours prévu en **février 2016**.
- Le **pic annuel** est clairement projeté en **décembre 2016**, légèrement supérieur au pic historique de décembre 2015.

---

### Prévisions par Lissage Exponentiel (Holt-Winters)

Nous ajustons un modèle de Lissage Exponentiel de Holt-Winters (multiplicatif) sur la série brute. Ce modèle détermine des coefficients de lissage ( $\alpha$  pour le niveau,  $\beta$  pour la tendance et  $\gamma$  pour la saisonnalité) pour projeter les composantes dans le futur.



	Point Forecast «2015»	Lo 80 «2015»	Hi 80 «2015»	Lo 95 «2015»	Hi 95 «2015»
Jan 2016	97.03990	92.56962	101.51018	90.20320	103.87661
Feb 2016	76.82962	72.34518	81.31406	69.97127	83.68798
Mar 2016	99.81859	95.28337	104.35380	92.88258	106.75460
Apr 2016	97.63018	93.06829	102.19207	90.65338	104.60699
May 2016	80.62116	76.06514	85.17718	73.65333	87.58899
Jun 2016	100.80940	96.15565	105.46315	93.69210	107.92669
Jul 2016	93.88711	89.21987	98.55434	86.74918	101.02503
Aug 2016	91.02116	86.32461	95.71771	83.83841	98.20392
Sep 2016	95.68821	90.91841	100.45802	88.39343	102.98300
Oct 2016	111.30944	106.36981	116.24907	103.75493	118.86395

1-10 of 12 rows

	Point Forecast «2015»	Lo 80 «2015»	Hi 80 «2015»	Lo 95 «2015»	Hi 95 «2015»
Nov 2016	111.09894	106.09534	116.10254	103.44659	118.75129
Dec 2016	130.13824	111.08635	149.19013	101.00089	159.27559

Previous 1 2 Next

## Comparaison des Performances et Modèle Retenu

Nous comparons les indicateurs d'erreur de chaque modèle sur la période d'ajustement (2010-2015).

Modèle	RMSE (Root Mean Squared Error)	MAE (Mean Absolute Error)
Paramétrique (Linéaire)	3,124	(Non calculé)
Holt-Winters (Multiplicatif)	4,214	3,104

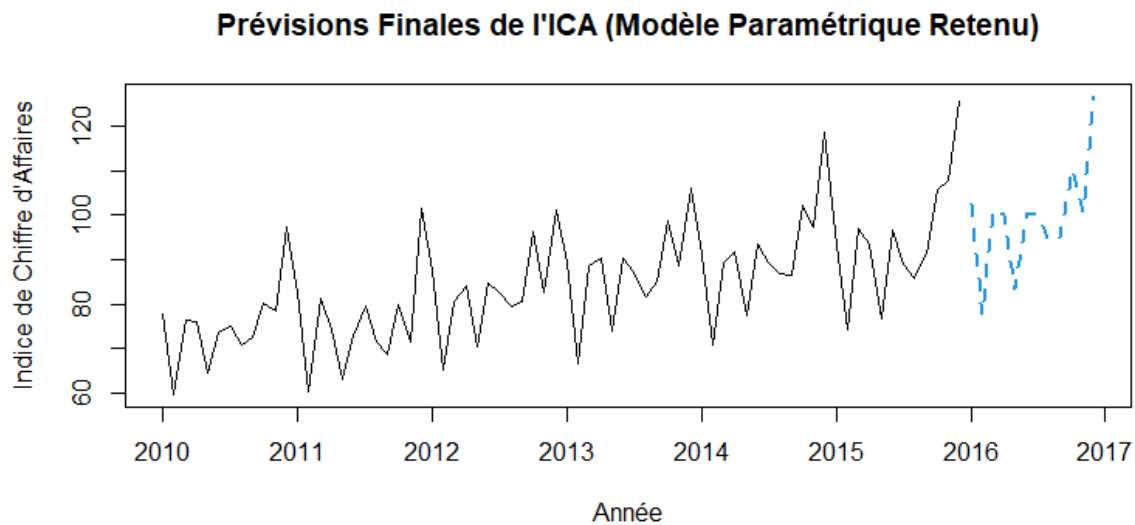
Décision : Le Modèle Paramétrique présente un RMSE de 3,124, ce qui est significativement inférieur au RMSE de 4,214 du modèle Holt-Winters.

Justification : Le Modèle Paramétrique est retenu comme le plus performant pour cette série temporelle. Ce résultat indique que la tendance de fond du marché de la location de logements est suffisamment linéaire et stable pour être mieux expliquée par une simple droite de régression que par l'approche plus dynamique du lissage exponentiel. Le modèle linéaire minimise l'erreur et est donc le plus fiable pour la projection à court terme.

## Prévisions Finales

Nous utilisons le Modèle Paramétrique retenu pour présenter les prévisions finales de l'ICA pour l'année 2016.





**Commentaire :** Le graphique final montre la projection du marché de la location pour 2016. Le modèle anticipe une **poursuite de la croissance stable** de l'ICA, avec un nouveau pic annuel en **décembre 2016** (approximativement 126,85).

## Conclusion :

En synthèse de cette analyse de série temporelle, l'Indice de Chiffre d'Affaires (ICA) du secteur de la location de logements (2010-2015) présente une **forte croissance** superposée à une **saisonnalité marquée et multiplicative**<sup>1</sup>.

1. **Modélisation :** Le modèle de décomposition multiplicatif a été retenu. La tendance de fond a été modélisée avec succès par une régression linéaire simple :  $\hat{T}_t = 71,40 + 0,356 \times t$ .
2. **Validation et Choix :** Le modèle paramétrique a été comparé à l'approche non paramétrique de Holt-Winters (HW). Le **Modèle Paramétrique** a démontré la meilleure performance sur l'échantillon d'apprentissage (RMSE = 3,124 contre 4,214 pour HW), confirmant que la tendance est suffisamment linéaire pour être bien capturée par une droite<sup>3</sup>.
3. **Prévision :** Le modèle retenu anticipe une **poursuite de la croissance** en 2016, avec une projection du pic annuel en décembre 2016 à environ 126,85.

