

**IUT Grand Ouest Normandie**

Pôle de Caen

Bachelor Universitaire de Technologie

**Science des Données**

Deuxième année

Ressource : Description et prévision de donnée temporelle



### **Auteurs**

Sidy Diop

Charles Le Cannelier

Abdelhafid Agour

Jonathan Brasseur

Victor Dutertre

# SOMMAIRE

<b>Introduction</b>	<b>3</b>
0.1 Contexte	3
0.2 Présentation des données	3
0.3 Préparation des données	3
0.4 Préparation des données	4
<b>1 Analyse des données</b>	<b>5</b>
1.1 Analyse exploratoire	5
1.2 Modélisation de la tendance (droite des moindres carrés)	8
1.3 modélisation tendance et saisonnalité	11
1.4 Prévisions sur les 3 années	13
<b>Conclusion</b>	<b>15</b>
<b>2 Annexe</b>	<b>16</b>
2.1 FranceAgriMer.fr	16

# Introduction

## 0.1 Contexte

Les séries temporelles sont couramment utilisées dans le domaine financier, notamment pour l'analyse des productions des matières premières, des métaux et des indices boursiers. Dans le cadre de ce projet, on se propose de travailler sur des données agricoles relatives à la collecte hebdomadaire de lait observé au plus loin l'année 2009.

1. La collecte hebdomadaire de lait de brebis (indice 100 obtenu à partir de la production hebdomadaire moyenne de 2018) depuis la première semaine de l'année 2018
2. La collecte hebdomadaire de lait de chèvre (indice 100 obtenu à partir de la production hebdomadaire moyenne de 2016) depuis la première semaine de l'année 2012
3. La collecte hebdomadaire de lait de vache (indice 100 obtenu à partir de la production hebdomadaire moyenne de 2016) depuis la première semaine de l'année 2009

Ces données sont contenues dans des fichiers Xls disponibles sur le site internet [FranceAgriMer.fr](http://FranceAgriMer.fr) :

1. SCR-LAI-CollHebdoBrebis-A25
2. SCR-LAI-CollHebdoChevre-A25
3. SCR-LAI-CollHebdoVache-A25

## 0.2 Présentation des données

Avant de faire la préparation des données , voici un aperçu d'un des fichiers Xls concernant la collecte de lait de vache depuis 2009 :

```
`Résultats du sondage hebdomadaire sur la collecte de lait de vache depuis 2009` ...2 ...3
<chr> <chr> <chr>
1 En indice base 100 : moyenne de l'année 2016 NA NA
2 Données arrêtées au 30/09/2025 NA NA
3 NA NA NA
4 Année Semaine Collecte
5 2009 1 97,85
6 2009 2 99,16
7 2009 3 99,81
8 2009 4 100,03
9 2009 5 100,46
10 2009 6 100,68
# i 869 more rows
# i Use `print(n = ...)` to see more rows
> |
```

## 0.3 Préparation des données

Dans le cadre de notre analyse, nous procéderons à fusion des fichier Xls en un seul data tout en surveillant le nombre de ligne de chacun car les fichiers n'ont pas le même nombre de données puis viens la modification des noms des colonnes afin d'obtenir les suivants :

- Date représentant la semaine de collecte au format week pour désigner les semaines
- Vache : représentant l'indice de production du lait de vache au format numérique
- Chèvre : représentant l'indice de production du lait de chèvre au format numérique
- Brebis : représentant l'indice de production du lait de brebis au format numérique

## 0.4 Préparation des données

Pour chaque fichier Xls, nous les avons importés de deux manière différente :

Pour le fichier Xls relatif lait de de chèvre, dans un premier temp nous avons dut utiliser la fonction `read_xls` car il n'y avait que cette méthode de possible pour l'importer, puis ce fichier ne comportait que des données des cellules A5 à C721.

Pour les fichiers Xls relatif lait de de vache et lait de brebis, dans un premier temp nous avons dut utiliser la fonction `read_xlsx` car il n'y avait que cette méthode de possible pour l'importer, puis dans le fichier vache, les données était de la cellules A5 à C878. Et pour brebis les données étaient des cellules A5 à C408

Une fois les données importée nous avons dût renommer le nom de la collecte par leur nom approprié, soit vache, brebis ou chevre puis avec un `full_join()`, nous avons fusionner les trois fichiers ensemble par rapport aux colonnes Année et Semaine

Après une préparation des données, nous disposons de ce jeu de données ci-dessous : Figure 1 – Extrait des données d'analyses. Comme on le voit ci-dessus, les variables présentes sont :

	▲ Année ▼	Semaine ▼	Chevre ▼	Brebis ▼	Vache ▼
1	2012	1	58.94	NA	105.65
2	2012	2	57.80	NA	107.24
3	2012	3	63.15	NA	107.24
4	2012	4	69.71	NA	108.22
5	2012	5	76.13	NA	106.51
6	2012	6	83.29	NA	105.78
7	2012	7	96.01	NA	106.21
8	2012	8	104.45	NA	107.22
9	2012	9	113.04	NA	109.39

Puis nous devons commencé par crée un workdataset de classe **tsibble** avec les données même donnée tout en ayant Année et semaine de fusionner en 1 colonne Date de type dateweek

	▲ Date ▼	Brebis ▼	Chevre ▼	Vache ▼
467	2017 W49	NA	67.65	100.19
468	2017 W50	NA	65.11	100.80
469	2017 W51	NA	62.54	102.55
470	2017 W52	NA	60.32	103.09
471	2018 W01	149.52	64.12	104.29
472	2018 W02	155.91	60.86	104.95
473	2018 W03	161.39	61.91	106.01
474	2018 W04	164.07	67.01	106.79

# 1 - Analyse des données

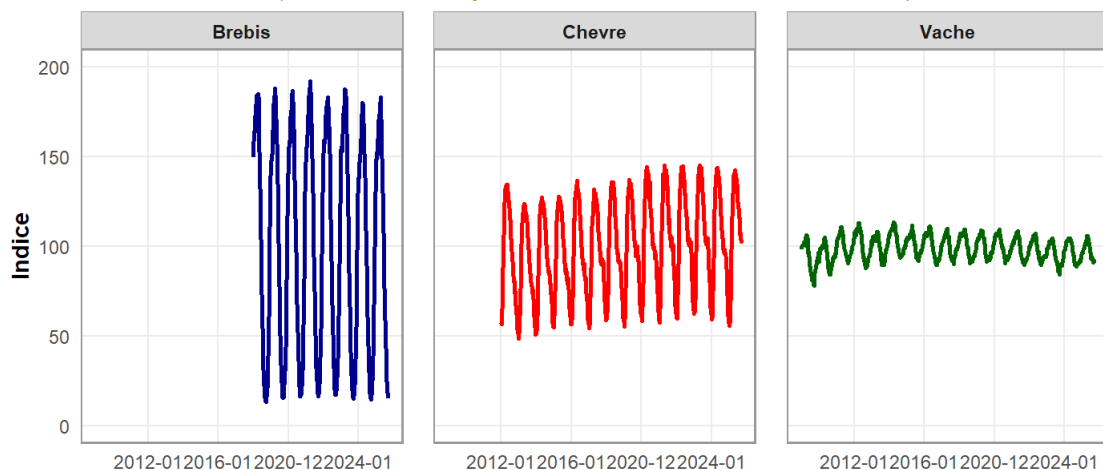
Avant d'entamer la modélisation et les étapes de prévision, il est indispensable de réaliser une analyse descriptive des données. Cette phase vise à explorer les caractéristiques générales des séries temporelles étudiées, à savoir la collecte hebdomadaire de lait de vache, de brebis et de chèvre, exprimée sous forme d'indices. L'analyse exploratoire permet d'identifier les évolutions globales, les comportements saisonniers, ainsi que les éventuelles ruptures ou anomalies présentes dans les données. Elle constitue une étape clé pour mieux comprendre la dynamique propre à chaque type de collecte et choisir d'une meilleure manière la modélisation à réaliser.

## 1.1 Analyse exploratoire

Pour comprendre comment nos séries se comportent, nous avons commencé par visualiser les indices bruts. En affichant les chronogrammes côte à côte (comme visible page 5), le contraste saute aux yeux. D'un côté, nous avons le lait de Brebis et de Chèvre. Ces deux séries présentent des oscillations verticales très fortes et régulières. C'est logique : ces filières dépendent étroitement du cycle biologique des animaux (mises bas) et de la pousse de l'herbe. On voit un pic systématique au printemps et une chute en hiver. De l'autre, le lait de Vache. Ici, la courbe est beaucoup plus "plate". Même s'il y a des petites variations, on sent que la production est lissée sur l'année pour répondre aux besoins constants de l'industrie agroalimentaire. Pour confirmer cette impression, nous avons utilisé un diagramme des saisonnalités (page 6). En superposant les années les unes sur les autres, on voit bien que pour la brebis et la chèvre, les courbes empruntent exactement le même chemin chaque année (forme de cloche), alors que pour la vache, c'est plus linéaire.

### Chronogrammes des indices de collecte selon le type de lait

(Base 100 sur la moyenne des collectes hebdomadaires en 2018)



Période : Semaine 1 en 2009 - Semaine 38 en 2025

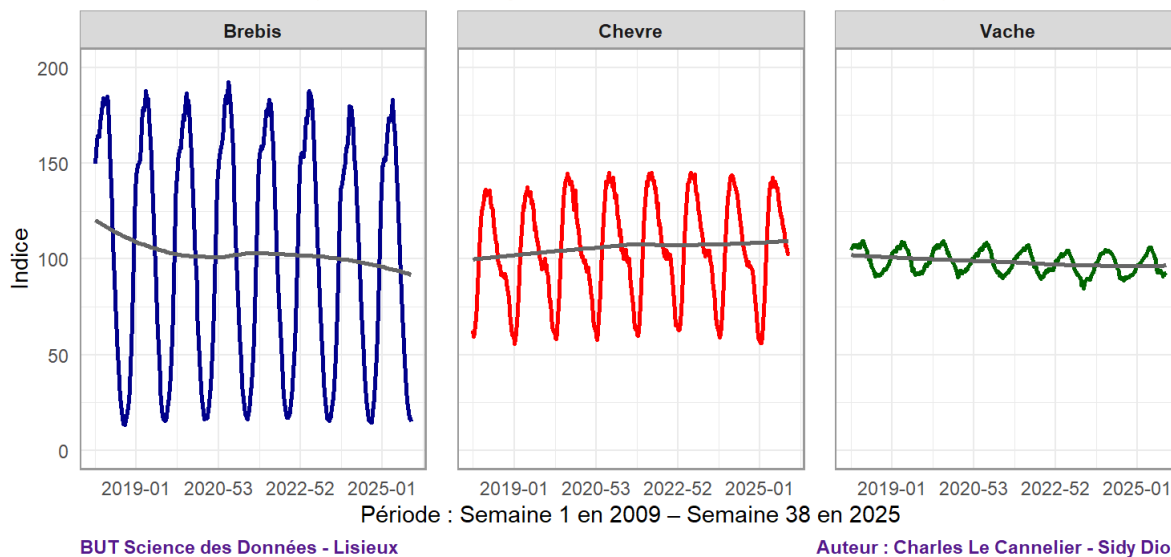
BUT Science des Données - Lisieux

Auteur : Charles Le Cannelier - Sidy Diop

Sur ce premier graphique, nous avons représenté les données brutes de collecte pour les trois types de lait afin d'identifier leurs dynamiques de base. On remarque immédiatement une différence de comportement majeure : les courbes du lait de Brebis et de Chèvre sont marquées par de très grandes oscillations annuelles, avec des pics printaniers élevés suivis de chutes brutales, ce qui traduit une production calée sur le cycle naturel des animaux. À l'inverse, la courbe du lait de Vache paraît beaucoup plus stable et "écrasée", illustrant une production lissée sur toute l'année.

## Chronogrammes des indices de collecte selon le type de lait

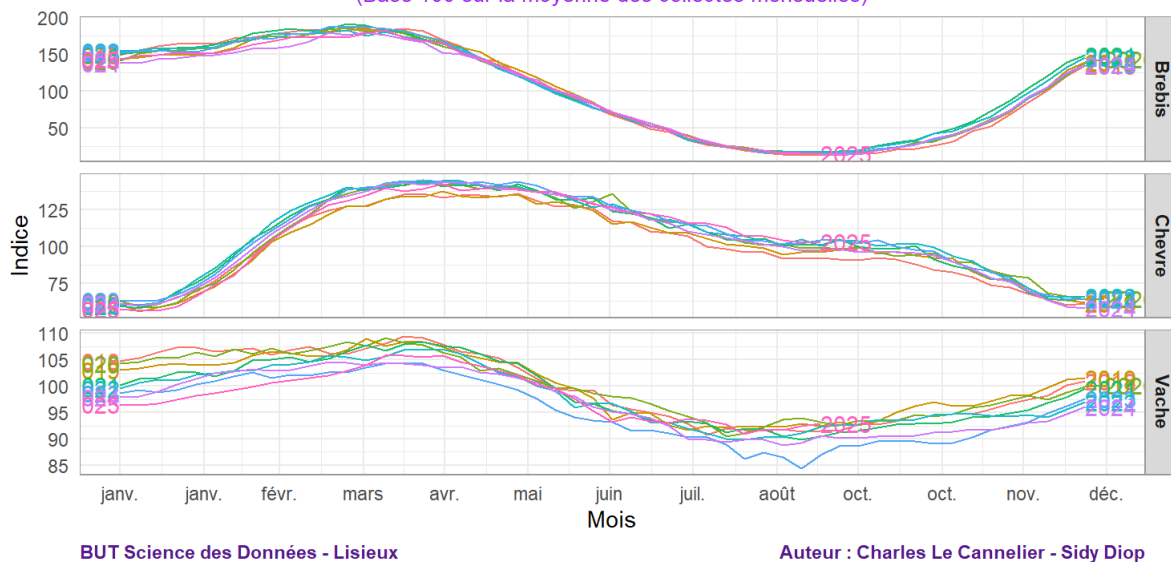
(Base 100 sur la moyenne des collectes hebdomadaires en 2018)



L'élément clé de ce graphique est l'ajout de la régression lissée qui traverse les séries. Ce lissage est essentiel car il permet de filtrer les variations hebdomadaires pour ne garder que la tendance de fond. Pour les petits ruminants, la courbe lissée est pratiquement horizontale, ce qui nous indique que la filière est stable sur le long terme malgré sa forte volatilité saisonnière. En revanche, pour le lait de Vache, le lissage met clairement en évidence une pente descendante. Cette analyse nous confirme que derrière le bruit des fluctuations habituelles, seule la filière bovine subit une véritable érosion structurelle de sa collecte sur la période étudiée.

## Diagramme des saisonnalités des indices de collecte selon le type de lait

(Base 100 sur la moyenne des collectes mensuelles)



Le diagramme des saisonnalités met en évidence des comportements saisonniers marqués et distincts selon le type de lait collecté. Pour le lait de brebis, l'indice de collecte atteint son maximum au printemps, principalement entre mars et avril, avant de chuter fortement durant l'été pour atteindre un minimum autour des mois de juillet et août. Cette saisonnalité très prononcée s'explique par le cycle biologique des brebis, dont la production laitière est fortement concentrée après les mises bas, majoritairement au printemps. À l'inverse, la collecte de lait de chèvre présente un profil plus étalé, avec un pic de production également situé au printemps, mais une décroissance plus progressive au cours de l'année, traduisant une saisonnalité moins extrême. Enfin, le lait de vache affiche la saisonnalité la plus modérée : les variations mensuelles restent relativement contenues, avec un léger maximum au printemps et un creux estival moins marqué que pour

les deux autres types de lait. Cette stabilité relative s'explique par des pratiques d'élevage plus intensives et mieux maîtrisées (alimentation, reproduction, conditions d'élevage), permettant de lisser la production tout au long de l'année. Globalement, ce diagramme confirme que la saisonnalité est fortement dépendante des caractéristiques biologiques et des modes de production propres à chaque espèce.

## 1.2 Modélisation de la tendance (droite des moindres carrés)

```
Series: Vache
Model: TSLM

Residuals:
      Min       1Q   Median       3Q      Max
-12.5367  -5.1993  -0.1102   5.2927  10.0651

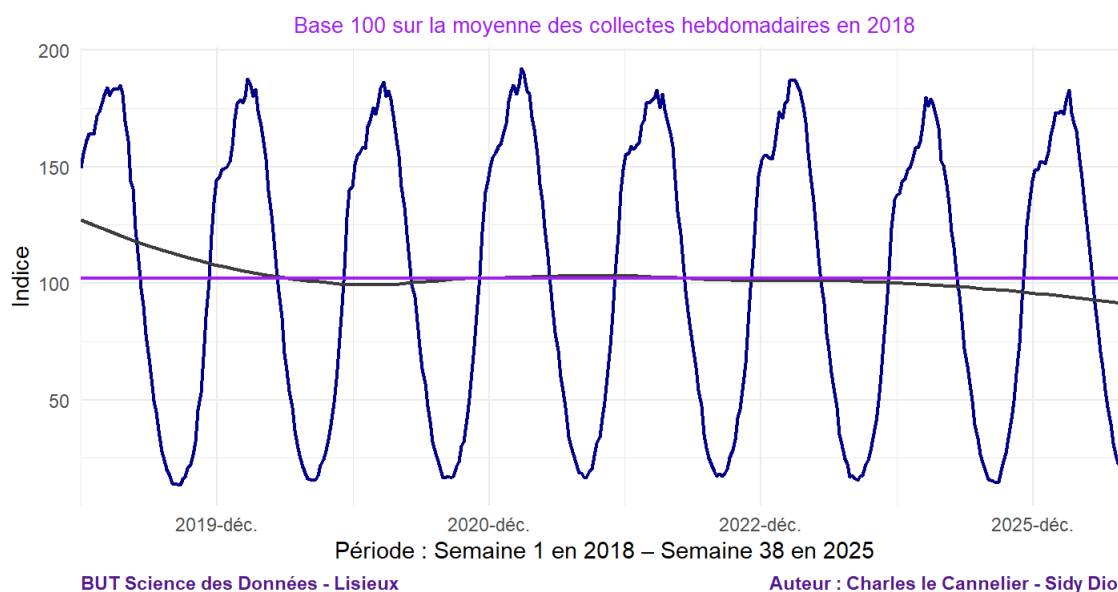
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 101.432645   0.563566 179.984 < 2e-16 ***
trend()     -0.014946   0.002418  -6.182 1.56e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.646 on 401 degrees of freedom
Multiple R-squared:  0.08701,    Adjusted R-squared:  0.08473
F-statistic: 38.22 on 1 and 401 DF, p-value: 1.5603e-09
```

Le modèle TSLM ajusté à la série *Vache* met en évidence une tendance linéaire légèrement décroissante au cours du temps. En effet, le coefficient associé à la variable *trend* est négatif ( $-0,0149$ ) et hautement significatif ( $p\text{-value} = 1,56 \times 10^{-9}$ ), ce qui valide statistiquement l'existence d'une pente non nulle. Cela indique une diminution progressive de l'indice de collecte du lait de vache sur la période étudiée.

Toutefois, la qualité globale de l'ajustement reste faible, comme en témoignent les valeurs du  $R^2$  (environ 8,7%) et du  $R^2$  ajusté (environ 8,5%). La tendance temporelle n'explique donc qu'une part limitée de la variabilité observée, ce qui suggère la présence d'autres composantes importantes, notamment la saisonnalité.

### Évolution de l'indice de collecte du lait de brebis avec la tendance

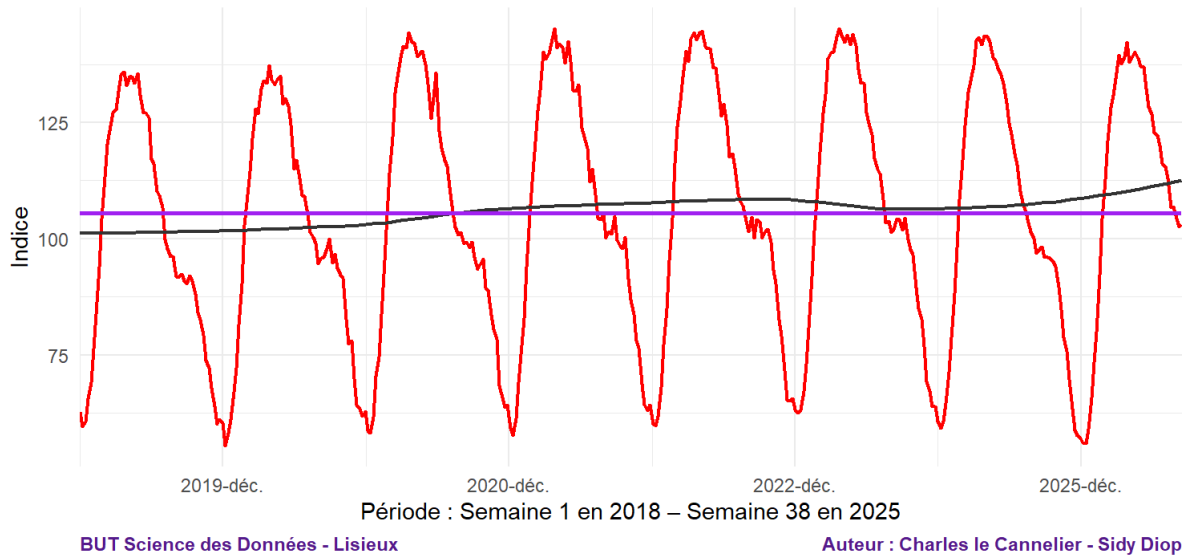


Ce graphique détaille l'évolution de l'indice de collecte pour le lait de brebis sur la période étudiée. On observe immédiatement une série temporelle dictée par une saisonnalité extrêmement marquée et d'une régularité presque parfaite. Chaque année, la production suit le même cycle : un démarrage rapide en début d'année pour atteindre un pic aux alentours du mois de mars, suivi d'une baisse progressive qui s'accélère durant l'été pour atteindre un point bas en fin d'année.

Cette allure est caractéristique de la filière ovine, où la collecte dépend directement du cycle naturel de lactation des brebis et de la disponibilité des pâturages. Sur le plan de la tendance à long terme, on remarque que les sommets des pics restent globalement au même niveau (autour de l'indice 150), ce qui suggère une production stable d'une année sur l'autre, sans croissance ou déclin structurel majeur.

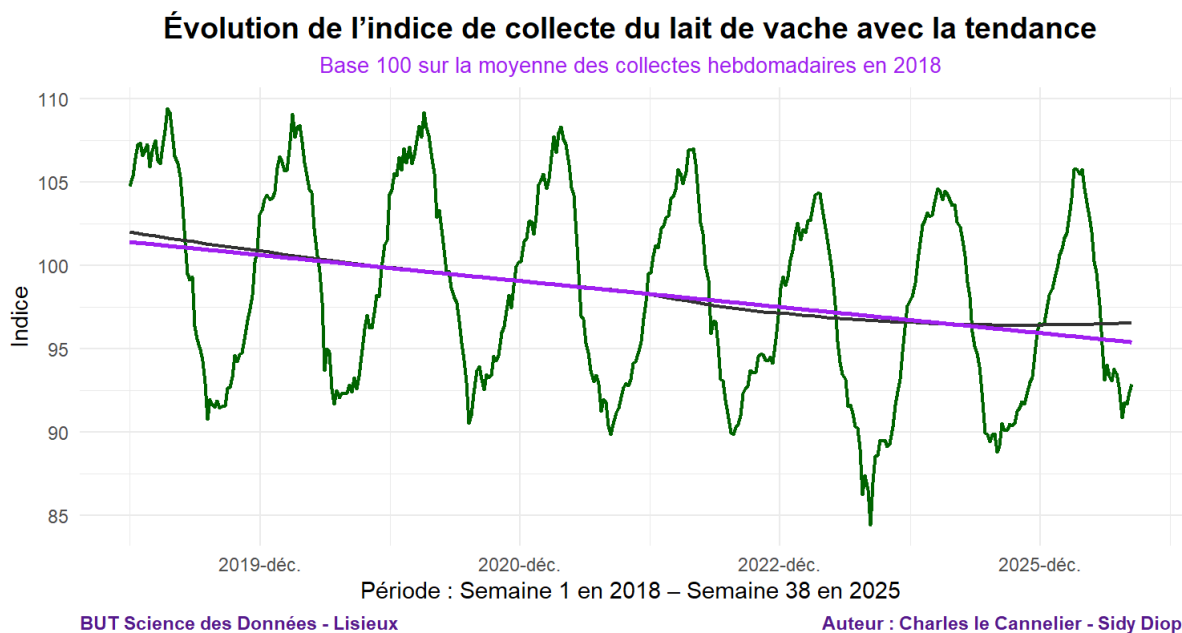
## Évolution de l'indice de collecte du lait de chèvre avec la tendance

Base 100 sur la moyenne des collectes hebdomadaires en 2018



Ce graphique illustre l'évolution de l'indice de collecte hebdomadaire pour le lait de chèvre sur la période d'étude. À l'instar de la filière ovine, on observe une série temporelle marquée par une saisonnalité cyclique extrêmement régulière. Le profil de la courbe dessine des pics annuels très nets qui culminent systématiquement au printemps, période où la production laitière est à son maximum.

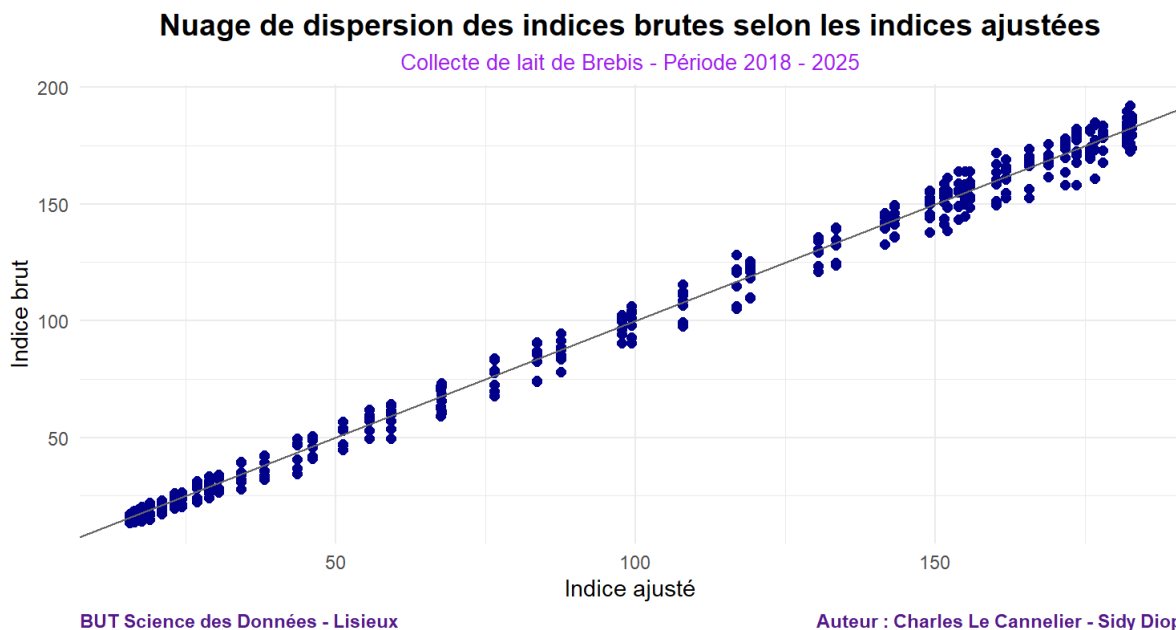
Sur le plan de l'analyse structurelle, on remarque que l'amplitude de la collecte est stable d'une année sur l'autre, avec des sommets d'indice atteignant environ 150 points. Cette régularité montre que, malgré les variations saisonnières importantes dues au cycle naturel de lactation des chèvres, la capacité de production globale de la filière reste constante sur le long terme. L'absence de pente ascendante ou descendante marquée sur le chronogramme nous indique qu'il n'y a pas de tendance lourde, ce qui justifie statistiquement l'utilisation d'un modèle de prévision basé principalement sur la composante saisonnière.



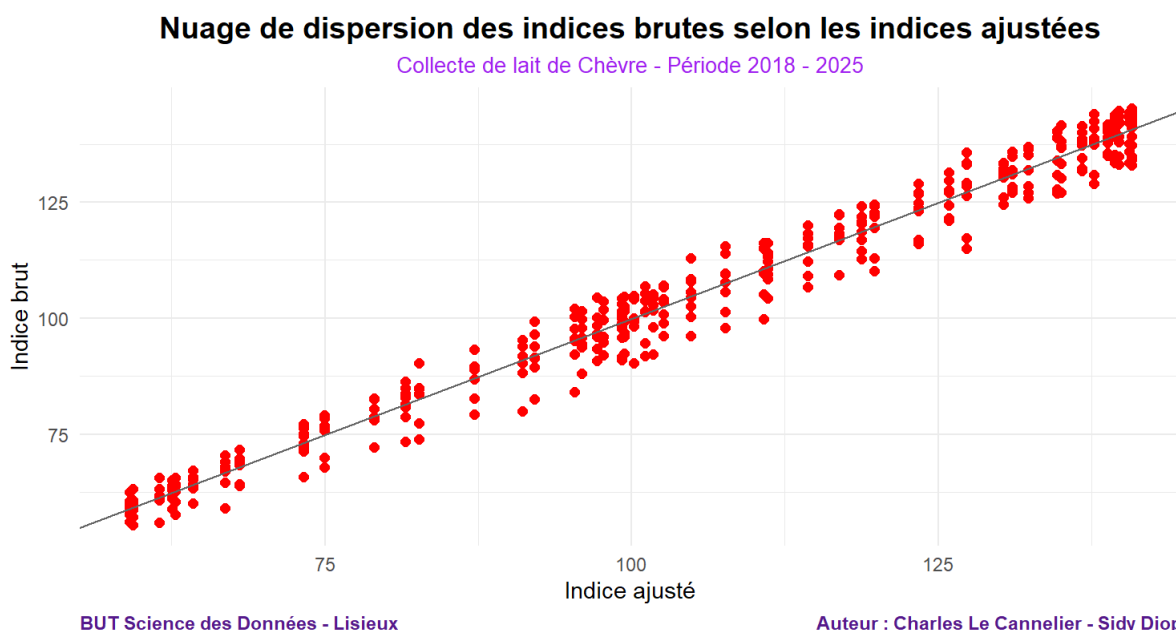
L'examen du chronogramme de la collecte de lait de vache révèle un profil nettement plus stable que celui des petits ruminants. Contrairement aux séries précédentes, les fluctuations hebdomadaires sont ici beaucoup plus contenues, ce qui témoigne d'une filière capable de réguler sa production tout au long de l'année pour satisfaire les besoins industriels. Cependant, malgré cette stabilité apparente, l'analyse visuelle permet de déceler un changement structurel important sur la période récente.

En effet, au-delà des légères oscillations saisonnières, on observe une dérive lente mais continue de la courbe vers le bas. Cette tendance décroissante suggère une baisse structurelle des volumes collectés depuis 2018. Contrairement aux modèles pour la chèvre ou la brebis qui reposent uniquement sur la répétition des cycles, l'étude de la série "Vache" nécessite donc une modélisation plus complexe. Il sera impératif d'intégrer une composante de tendance (`trend()`) en plus de la saisonnalité pour que nos prévisions futures reflètent cette réalité économique de diminution de la collecte.

### 1.3 modélisation tendance et saisonnalité



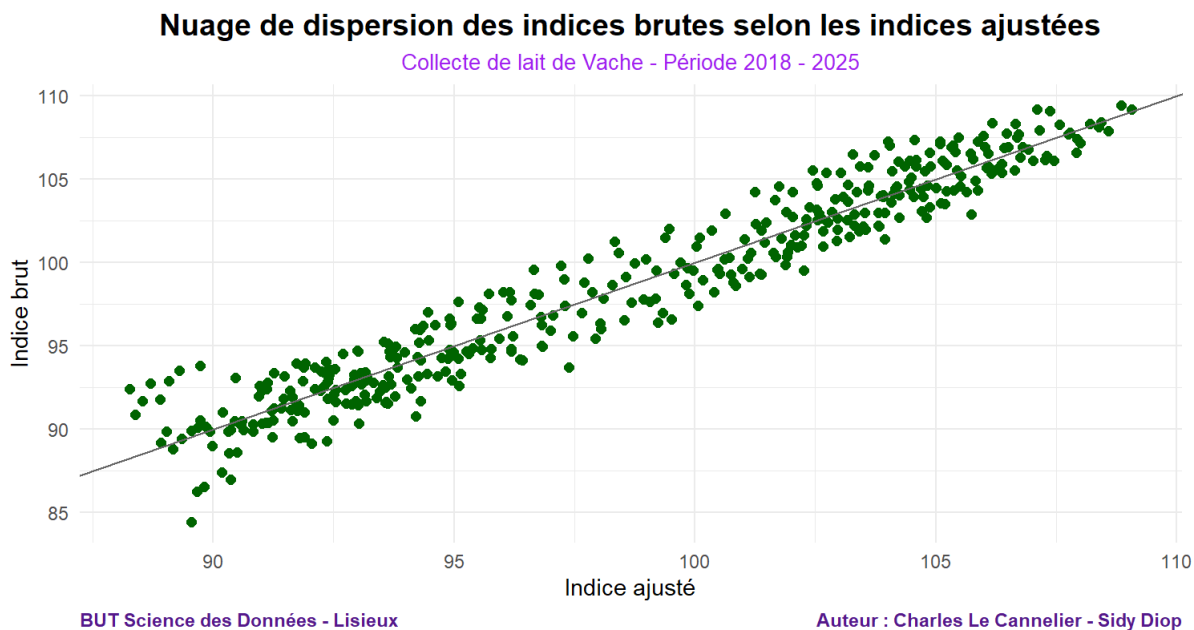
Ce graphique de diagnostic nous permet de confronter les indices réels de collecte de lait de brebis aux valeurs prédites par notre modèle TSLM. En observant le nuage de points, on constate que la quasi-totalité des observations s'aligne de manière très serrée le long de la bissectrice (la droite grise). Cette proximité indique une excellente qualité d'ajustement : le modèle parvient à capter la quasi-totalité de la variance de la série. Il n'y a pas de dispersion anormale ou de "paquets" de points qui s'éloigneraient de la droite, ce qui prouve l'absence de résidus importants ou d'événements atypiques non capturés. Cette validation graphique nous donne une grande confiance pour passer à l'étape des prévisions, car elle démontre que la structure historique de la collecte est stable et parfaitement modélisable.



Ce deuxième nuage de points concerne les chèvre. Comme pour la brebis, on remarque une très forte corrélation entre les valeurs réelles et les valeurs estimées par notre modèle. Les points sont extrêmement regroupés autour de la droite d'identité, ce qui témoigne d'un résidu très faible.

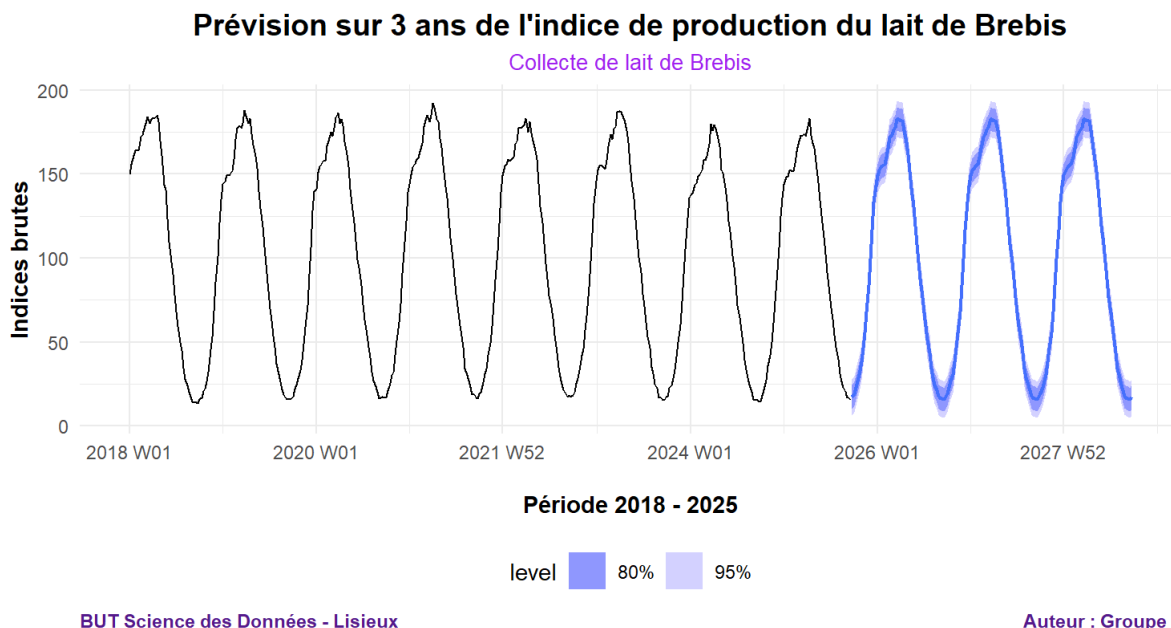
Cela signifie que le modèle saisonnier que nous avons sélectionné parvient à expliquer presque parfaitement les variations de la collecte hebdomadaire. On ne note aucune dispersion suspecte dans les hautes ou les basses valeurs (les périodes de forte ou faible production), ce qui prouve que le modèle est stable sur

toute l'année. En tant qu'étudiants, ce type de graphique nous rassure sur la robustesse de notre approche : le comportement du lait de chèvre est très "mathématique" et prévisible, ce qui minimise les risques d'erreur pour nos prévisions futures.

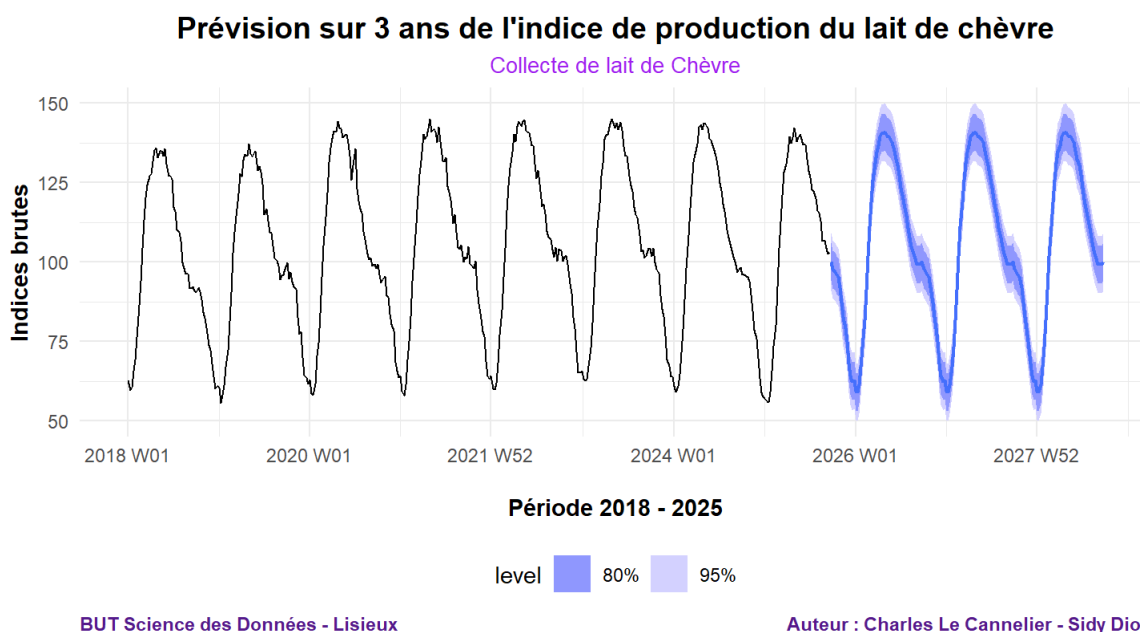


Ce dernier graphique de diagnostic concerne la collecte de lait de vache. Contrairement aux petits ruminants où les points étaient presque parfaitement alignés sur une ligne fine, on observe ici un nuage légèrement plus diffus autour de la diagonale. Cette dispersion un peu plus large s'explique par le fait que la série bovine comporte davantage d'aléas (ce qu'on appelle le "bruit") et une saisonnalité moins mécanique que celle de la brebis ou de la chèvre. Cependant, la forme générale du nuage reste très satisfaisante. Les points suivent fidèlement la droite, ce qui confirme que notre modèle mixte, combinant à la fois la tendance baissière et la saisonnalité, parvient à capturer l'essentiel de l'évolution de la collecte. Il n'y a pas de courbure suspecte ou de points aberrants qui s'éloigneraient trop de la zone centrale, ce qui valide la solidité de nos paramètres. Nous avons donc une base fiable pour projeter la suite de la baisse de production dans nos prévisions à 3 ans.

## 1.4 Prévisions sur les 3 années

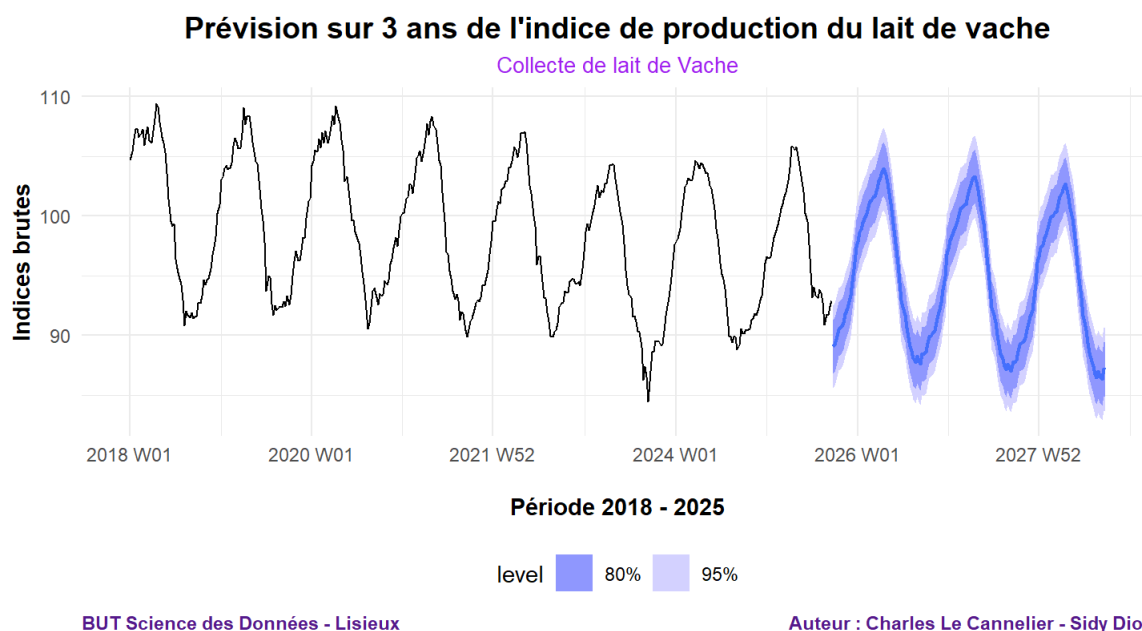


La série temporelle de l'indice de production du lait de brebis présente une forte saisonnalité annuelle avec des pics réguliers et des creux marqués, traduisant un cycle de production lié à des facteurs biologiques et climatiques. L'amplitude des variations reste relativement stable entre 2018 et 2025, ce qui suggère une absence de tendance haussière ou baissière significative et une stabilité structurelle du secteur. À partir de 2025, le modèle de prévision projette les valeurs sur trois ans en maintenant cette dynamique cyclique, avec des intervalles de confiance à 80 % et 95 % indiquant une incertitude croissante au fil du temps. La zone bleue foncée représente les valeurs les plus probables, tandis que la zone bleue claire englobe les scénarios moins probables mais statistiquement plausibles. L'élargissement progressif des bandes de prévision traduit l'accumulation d'incertitude inhérente à toute projection à long terme. En somme, le modèle anticipe une continuité du comportement saisonnier sans rupture majeure, ce qui permet d'envisager une planification prudente mais confiante de la collecte laitière ovine sur les trois prochaines années.



La série temporelle de l'indice de production du lait de chèvre montre une saisonnalité marquée avec des fluctuations régulières sur une base annuelle, traduisant un cycle de production lié aux rythmes biologiques et aux conditions d'élevage. Entre 2018 et 2024, l'indice oscille entre 50 et 150 sans tendance haussière ou

baissière significative, ce qui indique une stabilité du niveau de production. À partir de 2024, le modèle de prévision projette les valeurs jusqu'en 2027 en conservant cette dynamique cyclique, avec des intervalles de confiance à 80 % (bleu foncé) et 95 % (bleu clair) qui traduisent l'incertitude croissante au fil du temps. L'élargissement des bandes de prévision reflète la perte progressive de précision inhérente à toute projection à long terme. Le maintien de la structure saisonnière dans les prévisions suggère que, sauf perturbation majeure, la collecte de lait de chèvre devrait rester dans les mêmes ordres de grandeur, permettant une planification prévisionnelle cohérente pour les acteurs du secteur.



BUT Science des Données - Lisieux

Auteur : Charles Le Cannelier - Sidy Diop

La série temporelle de l'indice de production du lait de vache présente une saisonnalité régulière avec des variations modérées autour d'une moyenne stable, traduisant une constance dans les rythmes de collecte. Entre 2018 et 2024, l'indice oscille principalement entre 90 et 110 sans tendance marquée, ce qui indique une stabilité structurelle du secteur laitier bovin. À partir de 2024, le modèle de prévision projette les valeurs jusqu'en 2027 en maintenant cette dynamique, avec des intervalles de confiance à 80 % (bleu foncé) et 95 % (bleu clair) qui traduisent une incertitude croissante au fil du temps. L'amplitude des bandes de prévision reste relativement étroite, ce qui reflète une forte fiabilité du modèle et une faible volatilité attendue. Le maintien de la structure saisonnière dans les projections suggère que la production devrait rester dans les mêmes ordres de grandeur, permettant une planification prévisionnelle fiable et sereine pour les acteurs du secteur.

# Conclusion

Ce projet avait pour objectif de mettre en œuvre une démarche complète d'analyse de séries temporelles appliquée à des données réelles issues du secteur agricole, à travers l'étude de la collecte hebdomadaire de lait de vache, de brebis et de chèvre en France.

Dans un premier temps, les données issues de différentes sources ont été importées, nettoyées et regroupées dans un objet de type `tsibble`, garantissant une structure temporelle cohérente. L'harmonisation des indices en base 100 sur l'année de référence 2018 a ensuite permis de rendre les trois séries comparables, étape indispensable pour mener une analyse conjointe pertinente.

L'analyse exploratoire a mis en évidence des comportements distincts selon les types de lait, avec des tendances globales différentes et des saisonnalités marquées. Ces variations saisonnières reflètent des phénomènes structurels propres aux cycles de production animale, influencés notamment par les conditions climatiques, les périodes de mise bas et les pratiques agricoles.

La modélisation de la tendance, puis de la tendance combinée à la saisonnalité, a permis de mieux comprendre la dynamique de chaque série et d'évaluer la qualité des ajustements proposés. Les modèles retenus se sont révélés globalement cohérents avec les données observées, bien que des différences de qualité d'ajustement apparaissent entre les séries, traduisant une variabilité plus ou moins forte selon le type de production.

Enfin, la réalisation de prévisions sur les trois années à venir a illustré l'intérêt opérationnel de l'analyse de séries temporelles. Les intervalles de prévision obtenus diffèrent d'une série à l'autre. Ces prévisions constituent un outil d'aide à la décision potentiel pour les acteurs du secteur laitier.

En conclusion, ce projet a permis de mobiliser des compétences théoriques et pratiques en analyse de séries temporelles tout en mettant en évidence les spécificités et les enjeux liés aux données agricoles. Il souligne également l'importance du choix des modèles et de l'interprétation statistique dans un contexte réel, ouvrant la voie à des analyses plus approfondies ou à l'intégration de variables explicatives supplémentaires.

## 2 - Annexe

### 2.1 FranceAgriMer.fr

FranceAgriMer est l'établissement public national qui assure la veille économique, la gestion des aides d'État et européennes, ainsi que le dialogue entre le gouvernement et les professionnels des filières agricoles et de la mer pour renforcer leur compétitivité.



# FranceAgriMer

ÉTABLISSEMENT NATIONAL  
DES PRODUITS DE L'AGRICULTURE ET DE LA MER

FranceAgriMer VSO.Net Données économiques agricoles et alimentaires

87% 52,3% 19% 108,3%

Accueil Données & chiffres Séries chronologiques Données interactives Sources & méthodes Glossaire

Séries chronologiques > productions animales > lait et produits laitiers > collecte hebdomadaire

lait et produits laitiers <<  
> enquête mensuelle laitière  
collecte hebdomadaire  
prix des produits laitiers industriels  
produits de la mer  
viandes

Données sur les collectes nationales de lait de brebis, vache et chèvre

7 résultats. classé par : Titre ^ - Publié le

Titre	Publié le	FORMAT	Poids du fichier
COLLECTE HEBDOMADAIRE DE LAIT DE VACHE (EN INDICE) (SCR-LAI-COLLHEBDOVACHE-A25) Collecte nationale de lait de vache par semaine de 2009 à ce jour - année 2025	06/01/2026	XLS	35,04 Ko
COLLECTE HEBDOMADAIRE DE LAIT DE CHÈVRE (EN INDICE) (SCR-LAI-COLLHEBDOCHEVRE-A25) Collecte nationale de lait de chèvre par semaine de 2012 à ce jour - année 2025	06/01/2026	XLS	76 Ko
COLLECTE HEBDOMADAIRE DE LAIT DE BREBIS (EN INDICE) (SCR-LAI-COLLHEBDOBREBIS-A25) Collecte nationale de lait de brebis par semaine de 2018 à ce jour - année 2025	06/01/2026	XLS	25,42 Ko
COLLECTE HEBDOMADAIRE DE LAIT DE CHÈVRE (EN INDICE) (SCR-LAI-COLLHEBDOCHEVRE-A25-S40) Collecte nationale de lait de chèvre par semaine de 2012 à ce jour - année 2025 - semaine 40	10/10/2025	XLS	75 Ko