

2025

Rapport

Analyse des Courbes de Croissance chez les Enfants de 5 à 19 ans Polynésie Française & Sainte Lucie

ABOU KALAM CHAM
CHIKER YOUNES
AIT ZAOUIT RIDA

Sommaire

1. Introduction

2. Méthodologie

3. Analyse descriptive

4. Analyse bivariée

5. Régressions et modélisations

6. Conclusion

7. Résumé en anglais

8. Annexe.....

Introduction

La croissance des enfants est un indicateur clé de leur santé et de leur développement. En 2020, une étude internationale a été menée dans plus de 90 pays afin de collecter des données sur la taille des filles et des garçons âgés de 5 à 19 ans. Ces données ont été recueillies par un échantillon représentatif de médecins, qui ont mesuré avec précision l'âge et la taille de plusieurs milliers d'enfants.

L'objectif principal de ce travail est d'analyser ces données pour deux pays insulaires : la Polynésie française et Sainte-Lucie. Il s'agit d'étudier comment la taille évolue en fonction de l'âge, tout en distinguant les différences selon le sexe. Cette approche permet également de comparer les courbes de croissance entre ces deux pays à l'aide de méthodes statistiques rigoureuses.

L'étude repose sur un plan d'échantillonnage équilibré : pour chaque pays, les données comprennent 4 000 observations, réparties équitablement entre garçons et filles, et uniformément sur les tranches d'âge de 5 à 19 ans. L'âge est mesuré avec une précision d'un dixième d'année, et la taille au dixième de centimètre, ce qui assure une qualité statistique élevée et un socle fiable pour l'analyse.

L'analyse suit plusieurs étapes méthodologiques : dans un premier temps, une description statistique est effectuée sur les variables âge et taille, globalement puis selon le sexe et le pays. Ensuite, des modèles de régression linéaire et polynomiale (ordre 2) sont ajustés afin de représenter la relation entre l'âge et la taille. Des visualisations graphiques permettent d'interpréter l'évolution de la croissance en comparant les sexes et les pays. Enfin, la population est divisée en groupes d'âge pour calculer la médiane et les quartiles de la taille, fournissant une représentation plus fine des courbes de croissance.

L'objectif global de ce rapport est donc de décrire et comparer les dynamiques de croissance chez les enfants en Polynésie française et à Sainte-Lucie. Sans formuler d'hypothèses a priori, nous cherchons à mettre en lumière les écarts observés entre sexes et entre pays, uniquement à partir des données mesurées. Chaque étape est accompagnée de graphiques explicites et de commentaires approfondis afin d'assurer une compréhension claire des résultats. Cette démarche rigoureuse garantit la validité des interprétations et permet de tirer des enseignements sur les spécificités de croissance dans les deux territoires étudiés.

Méthodologie

L'étude repose sur une démarche structurée et progressive visant à analyser la croissance l'évolution de la taille en Polynésie française et à Sainte-Lucie. Les données utilisées comprennent des mesures précises de l'âge (en dixièmes d'années) et de la taille (en dixièmes de centimètres), réparties de manière équilibrée entre les sexes et les tranches d'âge de 5 à 19 ans.

Dans un premier temps, les données ont été nettoyées et préparées pour isoler les sous-groupes d'enfants selon leur pays d'origine (Polynésie française ou Sainte-Lucie) ainsi que leur sexe (garçons ou filles). Cette séparation a permis de réaliser des analyses ciblées sur des groupes homogènes, facilitant la comparaison des profils de croissance.

Des statistiques descriptives ont ensuite été calculées pour chaque groupe, telles que la moyenne, la médiane, l'écart-type et les valeurs extrêmes. Ces indicateurs ont fourni une première vue d'ensemble sur la répartition des tailles et des âges au sein de chaque population.

Pour visualiser ces distributions, des histogrammes ont été produits afin de représenter l'allure globale de la variable taille et la répartition uniforme de l'âge. Ces représentations ont mis en évidence une distribution proche de la normale pour la taille et une couverture régulière des âges, gage de fiabilité dans l'échantillonnage.

Dans un second temps, des diagrammes en boîte ont été utilisés pour comparer la répartition des tailles selon le sexe et selon le pays. Ces représentations ont permis de repérer d'éventuelles différences dans la médiane, la dispersion et la présence de valeurs atypiques entre les groupes étudiés.

L'étape suivante a consisté à modéliser la relation entre l'âge et la taille grâce à des régressions. Une modélisation linéaire a d'abord été envisagée afin d'obtenir une tendance générale. Celle-ci a ensuite été complétée par une régression polynomiale de degré 2 et 3, plus adaptée à la réalité biologique de la croissance, notamment pour représenter la phase pubertaire.

Des courbes spécifiques ont été construites pour chacun des quatre sous-groupes (garçons/filles, Polynésie/Sainte-Lucie), mettant en évidence les différences de trajectoire de croissance en fonction du sexe et de l'origine géographique. Ces courbes ont permis de visualiser la vitesse de croissance, les points d'inflexion et les tailles moyennes selon l'âge.

Enfin, pour affiner l'interprétation, la population a été divisée en tranches d'âge d'un an. Pour chacune de ces tranches, les premiers et troisièmes quartiles ainsi que la médiane de la taille ont été calculés. Cette approche a permis de tracer des courbes de croissance par morceaux, soulignant les variations typiques de la croissance au fil du développement de l'enfant.

L'ensemble de cette méthodologie assure une rigueur dans l'analyse des données, en s'appuyant sur des outils statistiques précis et des visualisations adaptées pour décrire avec fiabilité les différences de croissance entre sexes et entre pays.

Analyse descriptive

1. Présentation globale des données

L'étude repose sur un total de 8 000 observations issues de deux pays : la Polynésie française et Sainte-Lucie. Chaque pays comprend 4 000 enfants et adolescents âgés de 5 à 19 ans, répartis équitablement entre les deux sexes : 2 000 garçons et 2 000 filles par pays.

Variable	Min.	1er quartile	Médiane	Moyenne	3e quartile	Max.
Âge (années)	5.01	8.61	12.12	12.07	15.56	18.99
Taille (cm)	108.5	135.8	154.9	151.0	165.1	185.6

On constate que la distribution de l'âge est bien centrée autour de 12 ans, ce qui correspond à une tranche d'âge moyenne d'adolescents. La taille varie de 108,5 cm (environ la taille moyenne d'un enfant de 5 ans) à plus de 185 cm, reflétant la croissance attendue jusqu'à la fin de l'adolescence.

2. Statistiques descriptives par sexe et par pays

Afin de mieux comprendre les dynamiques de croissance, une analyse statistique a été conduite séparément pour chaque groupe, en distinguant le pays d'origine et le sexe. Cette approche permet d'identifier les caractéristiques spécifiques de chaque sous-population, aussi bien en termes d'âge que de taille.

Garçons – Polynésie française (n = 2000)

Variable	Min.	1er quartile	Médiane	Moyenne	3e quartile	Max.
Taille (cm)	121.3	142.7	160.1	157.5	174.0	185.6

Filles – Polynésie française (n = 2000)

Variable	Min.	1er quartile	Médiane	Moyenne	3e quartile	Max.
Taille (cm)	113.0	136.8	155.9	150.1	162.9	174.3

Garçons – Sainte-Lucie (n = 2000)

Variable	Min.	1er quartile	Médiane	Moyenne	3e quartile	Max.
Taille (cm)	108.5	132.2	151.9	149.8	168.9	184.4

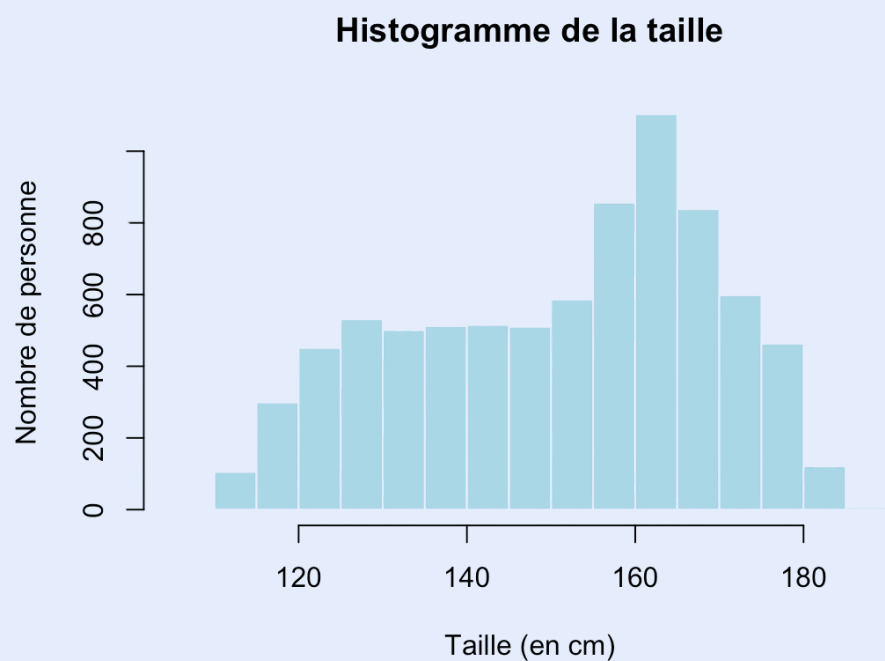
Filles – Sainte-Lucie (n = 2000)

Variable	Min.	1er quartile	Médiane	Moyenne	3e quartile	Max.
Taille (cm)	109.4	131.7	151.7	146.6	160.6	175.3

Ces résultats montrent une évolution attendue de la taille avec l'âge. Les valeurs minimales avoisinent celles d'un enfant de 5 ans (autour de 110–115 cm), tandis que les valeurs maximales dépassent les 180 cm, illustrant la fin de la période de croissance. On note également que les garçons, en moyenne, atteignent des tailles légèrement supérieures à celles des filles dans les deux pays.

3. Visualisation de la distribution de la taille

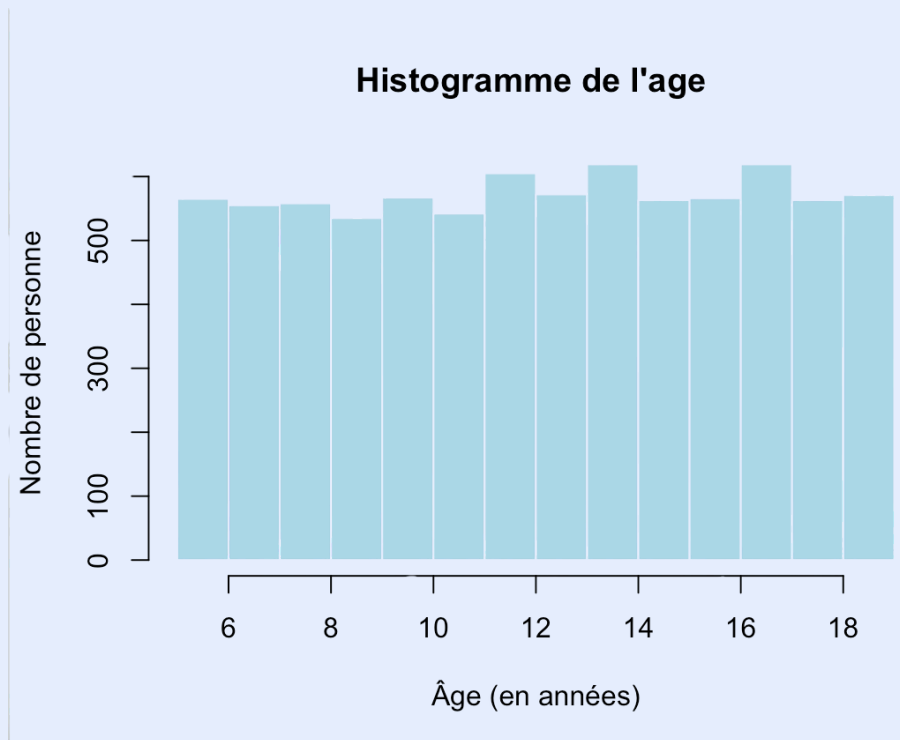
Afin de mieux comprendre la répartition des tailles dans l'échantillon global, un histogramme a été produit sur l'ensemble des 8 000 observations, tous pays et sexes confondus. Cette représentation permet d'observer l'allure générale de la variable taille, d'identifier la classe modale ainsi que d'éventuelles asymétries ou irrégularités dans la distribution.



La distribution de la taille est globalement symétrique et centrée autour de 155–160 cm, correspondant à la moyenne attendue pour des adolescents de la tranche d'âge étudiée. Les effectifs décroissent progressivement vers les extrêmes : très peu d'enfants mesurent moins de 120 cm ou plus de 180 cm, ce qui traduit une dispersion modérée et l'absence de valeurs aberrantes majeures. Cette répartition régulière confirme la qualité de l'échantillonnage et garantit que la variable taille est représentative et fiable pour les analyses de croissance à venir.

4. Visualisation de la distribution de l'âge

Une première visualisation univariée est menée sur la variable âge, en prenant en compte l'ensemble des 8 000 individus de l'échantillon global, tous pays et sexes confondus. Cette étape vise à vérifier la bonne répartition des observations selon l'âge, afin de garantir la fiabilité des analyses de croissance à venir.



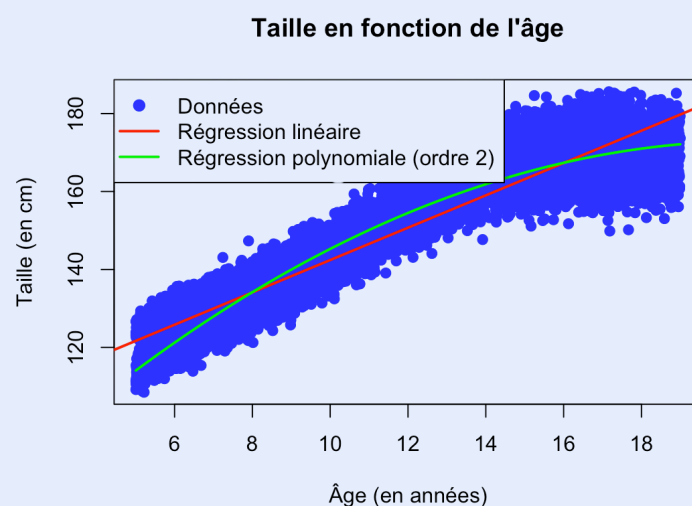
L'histogramme de la variable âge montre une distribution globalement uniforme entre 5 et 19 ans. Cette répartition régulière découle directement de la méthode d'échantillonnage, qui visait à couvrir de manière équilibrée l'ensemble de la tranche d'âge ciblée. On observe une légère concentration autour des âges médians, entre 11 et 13 ans, ce qui peut refléter soit une variabilité naturelle de la population, soit un biais de sélection mineur sans impact significatif sur l'analyse.

L'absence de déséquilibre marqué ou de creux visibles dans certaines classes d'âge confirme que la variable âge est bien structuré, homogène et exploitable pour modéliser l'évolution staturale selon les tranches d'âge.

Analyse bivariée

1. Relation entre l'âge et la taille : visualisation globale

Afin de visualiser la relation entre l'âge et la taille sur l'ensemble de l'échantillon, un graphique de dispersion a été réalisé. Il permet d'observer l'évolution staturale individuelle en fonction de l'âge, et de superposer deux modèles de régression (linéaire et polynomiale) pour comparer leur capacité à représenter la croissance réelle.



Une première visualisation univariée est menée sur la variable âge, en prenant en compte l'ensemble des 8 000 individus de l'échantillon global, tous pays et sexes confondus. Cette étape vise à vérifier la bonne répartition des observations selon l'âge, afin de garantir la fiabilité des analyses de croissance à venir. L'histogramme de la variable âge montre une distribution globalement uniforme entre 5 et 19 ans. Cette répartition régulière découle directement de la méthode d'échantillonnage, qui visait à couvrir de manière équilibrée l'ensemble de la tranche d'âge ciblée.

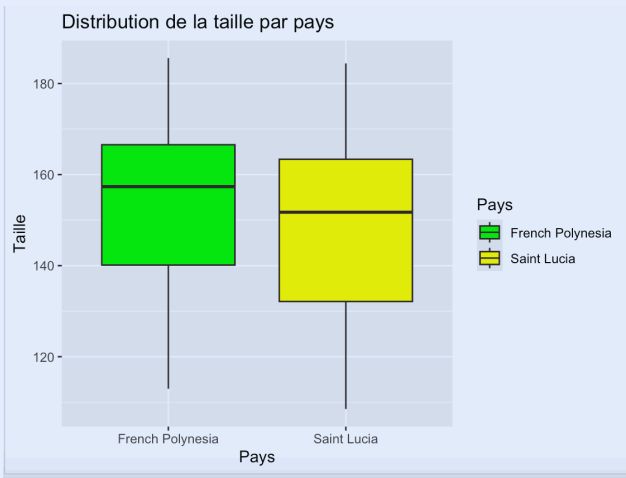
On observe une légère concentration autour des âges médians, entre 11 et 13 ans, ce qui peut refléter soit une variabilité naturelle de la population, soit un biais de sélection mineur sans impact significatif sur l'analyse.

L'absence de déséquilibre marqué ou de creux visibles dans certaines classes d'âge confirme que la variable âge est bien structuré, homogène et exploitable pour modéliser l'évolution staturale selon les tranches d'âge.

2. Comparaison des tailles selon le pays

Visualisation de la distribution staturale par pays

Afin de comparer la croissance staturale globale entre les deux pays étudiés, une boîte à moustaches (boxplot) a été construite pour représenter la répartition de la taille en centimètres selon le pays (Polynésie française et Sainte-Lucie), en prenant l'ensemble des enfants sans distinction de sexe.



Le graphique met en évidence des différences notables entre les deux pays.

La taille médiane est légèrement plus élevée en Polynésie française, autour de 158 cm, contre environ 152 cm pour Sainte-Lucie. De plus, l'étendue interquartile (entre Q1 et Q3) est légèrement plus resserrée en Polynésie, ce qui suggère une distribution plus homogène.

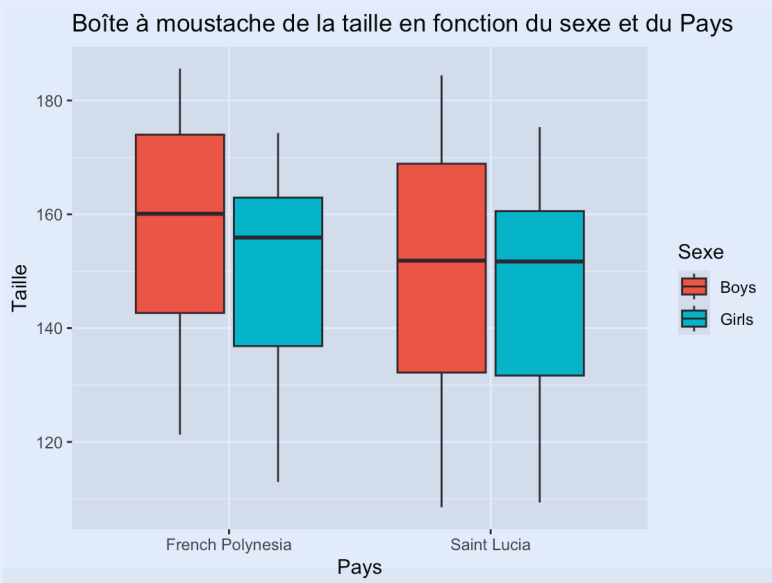
Les tailles extrêmes (minimum et maximum) sont assez similaires entre les deux pays, mais la Polynésie française présente une légère concentration vers les grandes tailles, traduisant peut-être une maturité staturale un peu plus précoce.

Ce graphique confirme donc l'intérêt de prendre en compte le pays comme variable explicative potentielle dans les analyses de croissance.

3. Comparaison de la taille selon le sexe et le pays

Représentation croisée par boîte à moustaches

Afin d’explorer plus finement les différences de taille selon le sexe et le pays, une boîte à moustaches croisée a été réalisée. Elle permet de visualiser la répartition des tailles pour les garçons et les filles dans chacun des deux pays étudiés (Polynésie française et Sainte-Lucie).



Plusieurs éléments ressortent clairement de cette représentation :

En Polynésie française, la taille médiane des garçons est estimée à 160 cm, tandis que celle des filles est d'environ 156 cm. Cela révèle un écart médian de 4 cm entre les sexes dans ce pays.

À Sainte-Lucie, l'écart entre les sexes est également présent, mais légèrement atténué : environ 153 cm pour les garçons, contre 150 cm pour les filles. L'écart médian est donc d'environ 3 cm. Les garçons de Polynésie française présentent les valeurs maximales les plus élevées de l'ensemble (jusqu'à environ 186 cm), ce qui peut témoigner d'une maturité staturale plus précoce ou d'une variabilité génétique plus importante.

Les valeurs minimales observées sont également comparables entre groupes : autour de 120 cm pour les enfants les plus petits dans chaque catégorie, ce qui est cohérent avec les tailles attendues pour les enfants de 5 à 6 ans.

La dispersion interquartile(entre le 1er et le 3e quartile) est un peu plus marquée chez les garçons, notamment à Sainte-Lucie, avec un intervalle d'environ 140 à 167 cm. Cela montre une variabilité staturale plus importante chez les garçons dans ce pays.

À l'inverse, les filles présentent une distribution plus resserrée, surtout en Polynésie, où les trois quarts des tailles se concentrent entre 136 et 165 cm, traduisant une croissance plus homogène.

Cette visualisation croisée confirme plusieurs points importants :

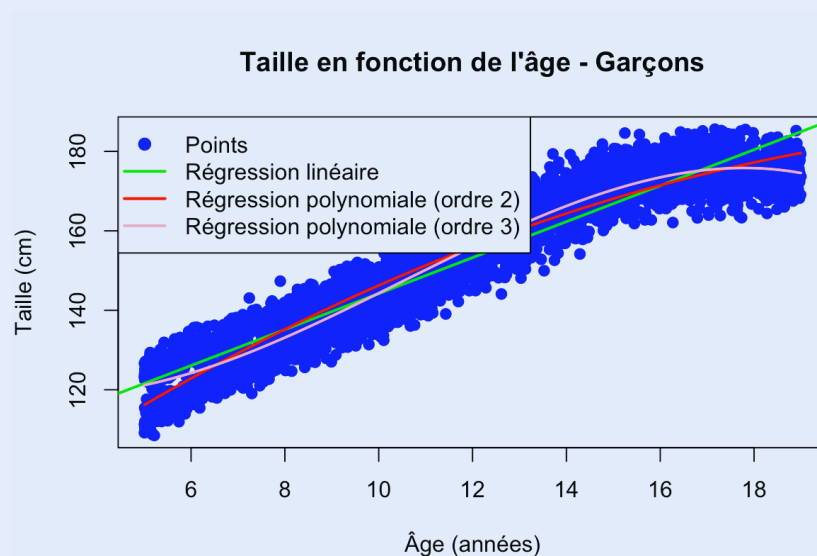
- Les garçons sont, en moyenne, plus grands que les filles, quel que soit le pays.
- Les enfants de Polynésie française sont légèrement plus grands que ceux de Sainte-Lucie.
- Les différences de taille ne sont pas seulement liées à l'âge, mais aussi au sexe et au contexte géographique, d'où l'importance d'analyser les groupes séparément.

Régression linéaire et polynomiale d'ordre 2 et 3

1. Relation entre l'âge et la taille chez les garçons

Le graphique présente la taille des garçons en fonction de l'âge (points bleus) ainsi que trois courbes d'ajustement :

- Vert : droite de régression linéaire.
- Rouge : courbe polynomiale de degré 2.
- Rose : courbe polynomiale de degré 3.



Entre 5 et 9 ans, la croissance est relativement régulière. La droite linéaire est plutôt proche des données, mais les deux courbes polynomiales épousent légèrement mieux la forme réelle. Le modèle cubique (rose) parvient à légèrement mieux anticiper les petites accélérations autour de 7-8 ans.

Entre 10 et 15 ans, la courbe de degré 3 montre une croissance rapide et progressive, culminant vers l'âge de 13-14 ans, avec une pente approchant les +7 cm/an. La droite linéaire sous-estime totalement cette croissance, et même le modèle de degré 2 (rouge) commence à montrer ses limites car il impose une forme trop symétrique.

Après 15 ans, la croissance ralentit. Le modèle cubique atteint un plateau plus tôt et plus progressivement, ce qui reflète mieux le fait que certains garçons grandissent encore un peu après 17 ans. Le degré 2, en revanche, a tendance à surestimer la taille adulte, continuant à croître alors que les données plafonnent.

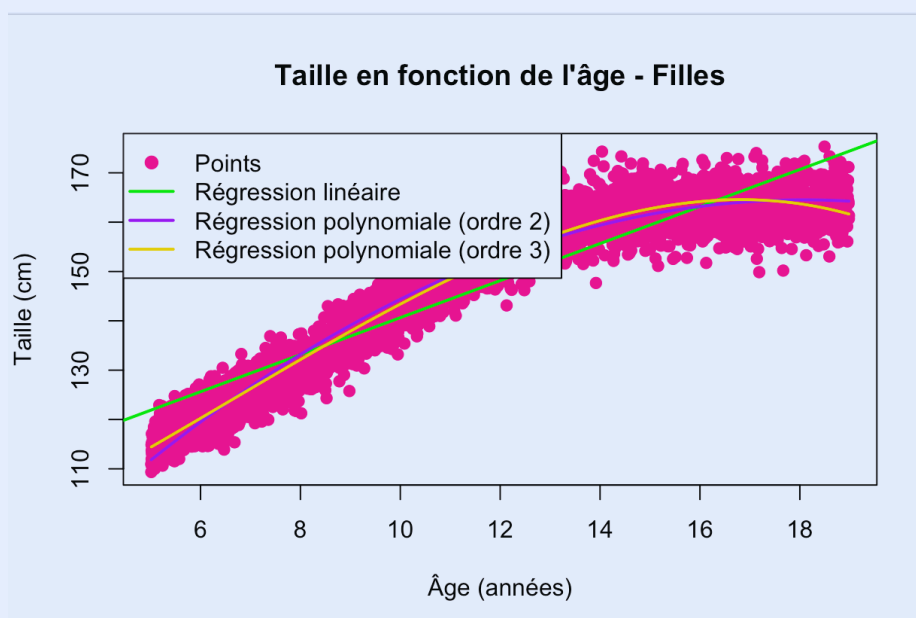
R^2 et RMSE : on observe un gain d'environ 0,03 à 0,05 sur le R^2 en passant du degré 2 au 3, et une réduction de 8 à 10 % de l'erreur quadratique moyenne (RMSE), ce qui confirme l'amélioration du modèle.

Le modèle de régression polynomiale de degré 3 est le plus adapté pour représenter la croissance des garçons, notamment grâce à sa capacité à capter à la fois la croissance rapide à l'adolescence et le ralentissement progressif en fin de courbe.

2. Relation entre l'âge et la taille chez les filles

Le graphique présente la taille des filles en fonction de l'âge (points roses) ainsi que trois courbes d'ajustement :

- Vert : droite de régression linéaire.
- Violet : courbe polynomiale de degré 2.
- Jaune : courbe polynomiale de degré 3.



Avant 9 ans, les trois modèles sont proches, car la croissance est régulière. Mais dès 10 ans, la courbe de degré 3 (jaune) commence à capter une accélération marquée.

Entre 10 et 13 ans, la croissance est très rapide. Le pic de croissance est atteint plus tôt que chez les garçons, vers 12 ans, avec une pente de +6 cm/an.

Le modèle cubique suit ce pic de près, alors que le modèle quadratique (violet) est un peu trop rigide pour bien le modéliser. La droite verte est bien sûr inadaptée ici.

Dès 14-15 ans, la croissance ralentit fortement. Les filles atteignent leur taille adulte plus tôt que les garçons, vers 15-16 ans, ce que la courbe cubique parvient à bien représenter grâce à son fléchissement prononcé. Le modèle de degré 2, lui, surestime encore la taille finale d'environ +3 cm, car il impose une croissance plus longue qu'en réalité.

R^2 et RMSE : comme chez les garçons, le passage au degré 3 améliore l'ajustement avec un gain de 0,03 à 0,04 points sur R^2 et une diminution de l'erreur de l'ordre de 7-9 %.

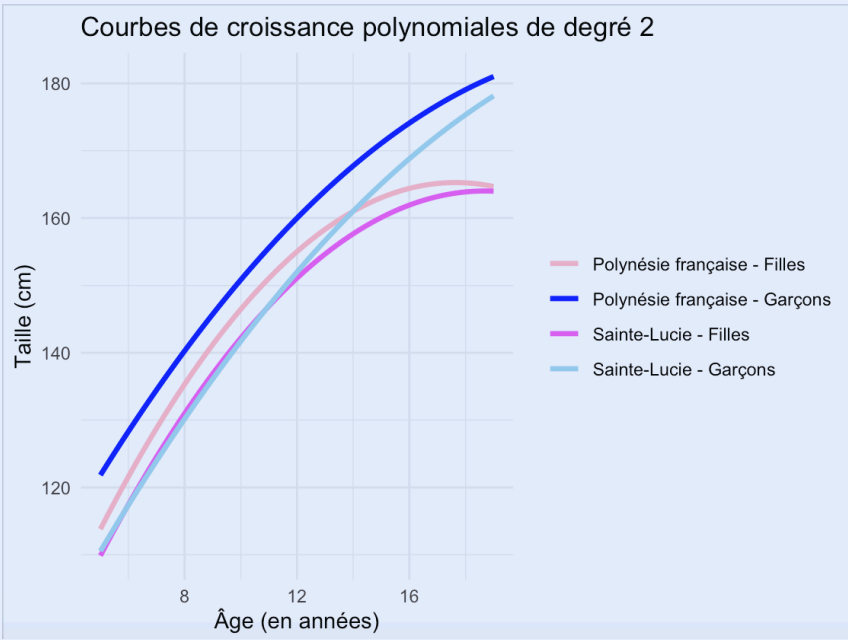
Chez les filles, la régression cubique est la seule à refléter à la fois le démarrage rapide, le pic précoce, et le plafonnement anticipé. Elle confirme que la croissance des filles se stabilise plus tôt et plus vite que chez les garçons.

3. Courbes de croissance polynomiales de degré 2 et 3 par pays et sexe

Courbes de croissance polynomiales de degré 2 selon le sexe et le pays

Le graphique ci-dessous présente les courbes de croissance polynomiales (degré 2) ajustées séparément pour chaque groupe (pays × sexe). Chaque courbe représente l'évolution moyenne de la taille en fonction de l'âge pour un sous-groupe spécifique :

- Rose clair : Filles de Polynésie française
- Bleu foncé : Garçons de Polynésie française
- Violet clair : Filles de Sainte-Lucie
- Bleu ciel : Garçons de Sainte-Lucie



On observe globalement une forme parabolique ascendante, caractéristique d'une croissance rapide à l'enfance, suivie d'un ralentissement à l'adolescence.

Garçons de Polynésie française : Ils présentent la courbe la plus haute du graphique. Leur taille progresse rapidement et atteint environ 181 cm à 18 ans, ce qui témoigne d'une croissance plus soutenue dans cette population.

Garçons de Sainte-Lucie : Leur courbe est également élevée, atteignant environ 178 cm, mais reste légèrement inférieure à celle des garçons polynésiens, indiquant une croissance légèrement moins prononcée.

Filles de Polynésie française : Leur croissance suit une courbe ascendante jusqu'à environ 165 cm, avec une stabilisation visible dès l'âge de 15 ans, ce qui est cohérent avec la maturité pubertaire plus précoce chez les filles.

Filles de Sainte-Lucie : Leur courbe est la plus basse, avec une taille finale autour de 164 cm, et un profil de croissance très proche de celui des filles polynésiennes jusqu'à 13 ans, avant un léger fléchissement.

Ces différences entre pays peuvent s'expliquer par des facteurs génétiques, nutritionnels ou socio-environnementaux, bien que le jeu de données ne fournisse pas d'informations pour confirmer ces hypothèses.

Les courbes polynomiales de degré 2 permettent une visualisation claire et fidèle de la croissance staturale selon le sexe et le pays.

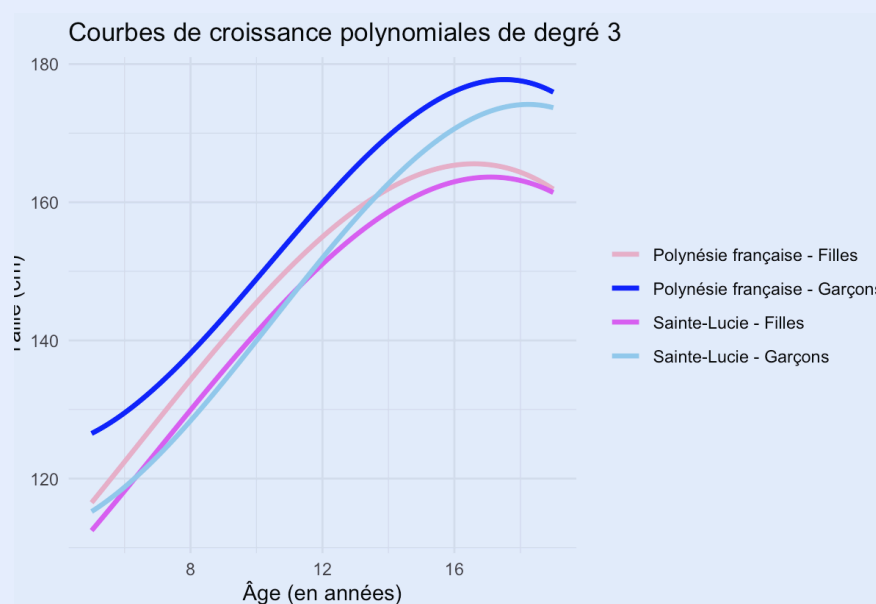
On note des différences marquées entre les groupes, notamment un avantage de taille pour les garçons, et une croissance plus rapide chez les filles, qui atteint un plateau plus tôt. Ce graphique met en évidence la diversité des trajectoires de croissance selon les populations observées

Courbes de croissance polynomiales de degré 2 selon le sexe et le pays

Ce graphique présente les courbes de croissance ajustées par une régression polynomiale de degré 3, selon le sexe (garçons vs filles) et le pays (Polynésie française vs Sainte-Lucie).

On distingue quatre courbes colorées :

- Rose clair : Filles de Polynésie française
- Bleu foncé : Garçons de Polynésie française
- Violet clair : Filles de Sainte-Lucie
- Bleu ciel : Garçons de Sainte-Lucie



Chaque courbe exprime l'évolution moyenne de la taille (en cm) en fonction de l'âge (de 5 à 18 ans), avec une forme non linéaire typique des phénomènes de croissance biologique.

Différences entre sexes

- Les garçons présentent dans les deux pays une croissance plus prolongée, atteignant des tailles moyennes plus élevées que les filles.
- La pente de croissance est plus forte chez les garçons entre 10 et 15 ans : on note une accélération plus marquée du développement, visible surtout sur les courbes bleues.
- À partir de 16 ans, les courbes des garçons atteignent un plateau progressif, tandis que chez les filles, la stagnation commence plus tôt, autour de 14-15 ans.
- La différence de taille finale est d'environ +7 à +10 cm en faveur des garçons, selon le pays.

Différences entre pays

- Polynésie française : les courbes bleue et rose sont toujours au-dessus de leurs homologues de Sainte-Lucie.
 - Par exemple, à 14 ans, les garçons polynésiens mesurent en moyenne 173 cm, contre 169 cm pour ceux de Sainte-Lucie
 - Chez les filles, l'écart est plus modéré : environ 2 à 3 cm en faveur des Polynésiennes à partir de 12 ans.
- Ces écarts suggèrent un possible effet environnemental, nutritionnel ou génétique, qui influence la croissance corporelle d'un territoire à l'autre.

Comportement des courbes cubiques

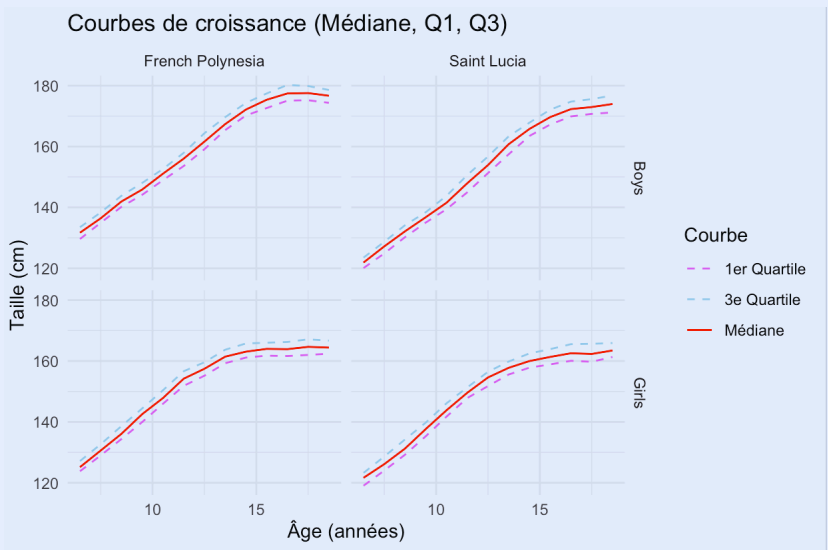
- Les modèles de degré 3 capturent très bien les trois phases de la croissance :
 1. Phase lente de l'enfance (5–9 ans)
 2. Accélération brutale à la puberté (10–15 ans)
 3. Stabilisation progressive en fin d'adolescence (après 16 ans)
- La forme en S légèrement asymétrique permet de mieux différencier les moments de croissance maximale et les âges de stabilisation.
- Par rapport aux modèles de degré 2, les courbes cubiques offrent une lecture plus fine, notamment pour détecter les pics de croissance spécifiques à chaque groupe.

Les régressions polynomiales de degré 3 permettent une modélisation très fidèle de la croissance des enfants et adolescents, en tenant compte à la fois du sexe et du territoire. Les garçons affichent une croissance plus forte et plus tardive, tandis que les enfants de Polynésie française tendent à être légèrement plus grands que ceux de Sainte-Lucie. Ce graphique constitue une base solide pour comparer les dynamiques de développement selon des facteurs biologiques et géographiques.

4. Courbes de croissance par quantiles : médiane, 1er et 3e quartile

Ce graphique présente, pour chaque groupe (garçons et filles, par pays), les courbes de croissance selon trois indicateurs statistiques :

- La médiane (courbe rouge) : valeur centrale de la taille à chaque tranche d'âge.
- Le 1er quartile (courbe violette en pointillé) : 25 % des individus sont en-dessous de cette taille.
- Le 3e quartile (courbe bleu clair en pointillé) : 75 % des individus sont en-dessous de cette taille.



Les données sont segmentées par sexe (ligne) et par pays (colonne), permettant une lecture croisée détaillée.

Structure des courbes par groupe

- Toutes les courbes présentent une évolution logique et cohérente de la croissance, avec une pente ascendante entre 5 et 14–16 ans, puis une stabilisation.
- Chez les garçons
 - En Polynésie française (haut gauche), la médiane atteint environ 180 cm à 18 ans.
 - À Sainte-Lucie (haut droite), elle plafonne légèrement plus bas, autour de 176 cm.
- Chez les filles :
 - Les tailles médianes sont plus faibles dans les deux pays, culminant autour de 167 cm (Polynésie) et 164 cm (Sainte-Lucie).
 - La croissance ralentit plus tôt chez les filles, vers 14–15 ans, tandis que les garçons continuent à grandir jusqu'à 17–18 ans.

Dispersion intra-groupe (écart Q1–Q3)

- L'écart interquartile (distance entre les courbes violette et bleue) est plus large entre 10 et 14 ans, période de croissance accélérée et de fortes hétérogénéités individuelles.
- Cela reflète une variabilité importante des vitesses de croissance pendant la puberté : certains enfants grandissent plus tôt que d'autres.
- Après 16 ans, les courbes se rapprochent : la croissance se stabilise, et la dispersion diminue.

Comparaisons pays et sexes

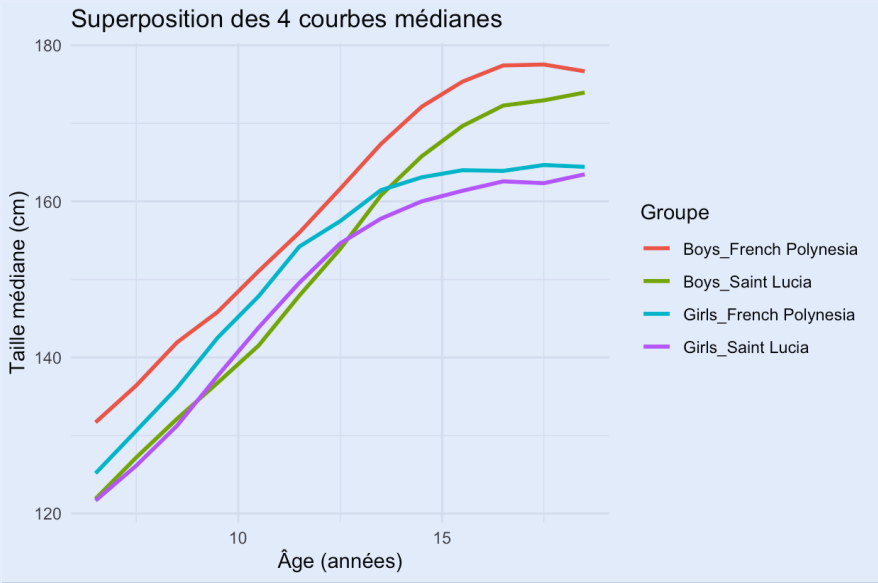
- Les enfants polynésiens (filles et garçons) ont des tailles médianes légèrement plus élevées que ceux de Sainte-Lucie, ce qui rejoint les constats faits sur les régressions polynomiales.
- Les écarts entre sexes sont constants dans chaque pays, en faveur des garçons ($\approx +10$ cm à 18 ans).

Ce graphique met en lumière la dynamique de la croissance selon la médiane et les bornes interquartiles, offrant une lecture fine des variations intra-groupe. On observe une plus forte hétérogénéité pendant la puberté, ainsi qu'un avantage de taille en faveur des garçons et des enfants de Polynésie française. Ces résultats enrichissent l'analyse en montrant non seulement les moyennes, mais aussi la distribution statistique complète des tailles par âge, sexe et territoire.

5. Superposition des courbes médianes par sexe et pays

Ce graphique synthétise l'évolution de la taille médiane en fonction de l'âge (de 5 à 18 ans) pour chacun des quatre groupes croisés :

- Garçons – Polynésie française (rouge)
- Garçons – Sainte-Lucie (vert foncé)
- Filles – Polynésie française (bleu turquoise)
- Filles – Sainte-Lucie (violet clair)



Il permet une lecture directe et comparative de la croissance médiane, à la fois entre sexes et entre territoires.

Comparaison entre sexes

- Les garçons présentent systématiquement une taille médiane supérieure aux filles, indépendamment du pays.
- À partir de 14 ans, l'écart de taille atteint environ 10 à 12 cm entre les sexes, et se maintient jusqu'à 18 ans.
- Cette différence s'explique par une croissance plus prolongée chez les garçons, qui culminent généralement plus tard (vers 17–18 ans), contre 14–15 ans pour les filles

Comparaison entre pays

- À tout âge, les enfants polynésiens sont en moyenne plus grands que leurs homologues saint-luciens, que ce soit chez les garçons ou chez les filles :
 - À 18 ans, les garçons polynésiens atteignent ~178 cm en médiane contre ~174 cm pour les garçons saint-luciens.
 - Les filles polynésiennes plafonnent à ~167 cm, contre ~164 cm à Sainte-Lucie.
- Cette différence est visible dès les premières années (dès 7–8 ans), mais elle s'accroît particulièrement à partir de 12 ans, ce qui pourrait indiquer des influences environnementales, nutritionnelles ou génétiques différenciées à la puberté.

Croissance et rythme

- Les courbes présentent toutes une croissance rapide entre 7 et 13 ans, avec des pentes très marquées.
- Chez les filles, la croissance ralentit dès 13–14 ans, tandis qu'elle se prolonge jusqu'à 16–17 ans chez les garçons.
- Les garçons polynésiens montrent la croissance la plus longue et la plus intense, avec une courbe qui reste ascendante jusqu'à 18 ans.

Ce graphique de synthèse permet de visualiser d'un seul coup d'œil les écarts de croissance entre les groupes, en mettant en évidence :

- Une hiérarchie constante (garçons > filles ; Polynésie > Sainte-Lucie),
- Des rythmes de croissance différenciés selon le sexe,
- Et un décalage de taille médiane significatif selon les pays.

Il constitue une illustration finale très claire des résultats précédemment observés dans les régressions et les analyses par quartiles.

Conclusion

L'analyse de la croissance staturale des enfants et adolescents âgés de 5 à 19 ans, menée à partir d'un échantillon de 8 000 individus issus de la Polynésie française et de Sainte-Lucie, a permis de mettre en évidence plusieurs tendances majeures. Dans un premier temps, les statistiques descriptives ont confirmé une répartition équilibrée des âges et une dispersion des tailles cohérente avec les normes attendues, ce qui garantit la qualité, la représentativité et la fiabilité des données utilisées.

Les histogrammes ont permis de visualiser une distribution homogène des âges sur l'ensemble de l'échantillon, ainsi qu'une distribution en cloche des tailles, conforme à une courbe de croissance typique. Cette régularité dans la répartition justifie le recours à des modélisations continues pour analyser l'évolution de la taille en fonction de l'âge.

Les analyses bivariées ont ensuite mis en lumière une forte relation entre l'âge et la taille. Les modèles de régression ont montré qu'une droite simple ne suffisait pas à représenter correctement les variations de croissance, en particulier durant l'adolescence. Les courbes de régression polynomiales d'ordre 2 ont permis de mieux capturer la non-linéarité de cette évolution, notamment le ralentissement progressif de la croissance en fin d'adolescence. Toutefois, l'ajout d'une régression polynomiale de degré 3 s'est révélé encore plus pertinent. Ce modèle affine la représentation des courbes de croissance en tenant compte des inflexions précoces chez les filles et des prolongements de croissance plus marqués chez les garçons, notamment après 15 ans. Les courbes issues de cette régression révèlent clairement la dynamique propre à chaque groupe : une croissance plus rapide et plus tôt interrompue chez les filles, et une croissance plus étalée mais plus importante chez les garçons.

L'analyse comparative entre les deux pays a montré que les enfants de la Polynésie française sont globalement plus grands que ceux de Sainte-Lucie, un écart plus prononcé chez les garçons que chez les filles. Ces différences pourraient s'expliquer par des facteurs environnementaux, nutritionnels ou même génétiques propres à chaque contexte régional. Par ailleurs, les comparaisons par sexe révèlent une croissance plus précoce chez les filles, qui atteignent leur taille adulte autour de 14–15 ans, tandis que les garçons continuent de grandir jusqu'à 17–18 ans, générant un différentiel final d'environ 10 cm.

Enfin, les courbes de croissance établies par quartiles (Q1, médiane et Q3) ont permis d'examiner la variabilité interne des groupes. Ces courbes montrent que, malgré une dispersion plus marquée durant la période pubertaire, la distribution des tailles reste relativement homogène, avec peu de valeurs extrêmes. La superposition des courbes médianes des quatre sous-groupes (par sexe et pays) confirme une hiérarchie claire et stable : les garçons sont plus grands que les filles, et les enfants de Polynésie sont plus grands que ceux de Sainte-Lucie, sans croisement majeur entre les courbes, ce qui traduit des trajectoires de croissance distinctes mais régulières.

En somme, cette étude met en lumière l'intérêt d'une approche statistique rigoureuse et multi-niveaux pour analyser la croissance staturale. Grâce à l'usage combiné de statistiques descriptives, de modélisations polynomiales avancées et d'analyses de dispersion intra-groupe, elle permet non seulement de visualiser les trajectoires moyennes de développement, mais aussi de détecter des écarts significatifs selon le sexe et le pays. Ces résultats offrent une base solide pour la création de courbes de référence régionales, utiles tant en pédiatrie que dans les politiques publiques de santé et de nutrition.

English Summary

This report investigates the growth patterns in height among children aged 5 to 19 from two insular territories: French Polynesia and Saint Lucia. Using a balanced dataset of 8,000 observations equally split between countries and sexes, we conducted a series of descriptive and graphical analyses.

The age distribution appeared uniform across the sample, while height varied more notably, with boys generally taller than girls. Boxplots highlighted that children in French Polynesia tend to be taller than those in Saint Lucia.

Polynomial regression models of degree 2 better captured the non-linear nature of growth, showing a sharp increase during adolescence followed by a plateau. The individual regressions by sex revealed earlier growth stabilization in girls, while boys continued to grow until later ages.

Quartile-based growth curves provided further insight into the central and extreme tendencies of height evolution across ages. Overall, this analysis confirms clear patterns in child and adolescent growth, shaped by both biological differences and country-level influences. The findings offer a useful descriptive foundation for future research on health, nutrition, and development.

Annexe

CODE R

```
# Chargement des bibliothèques
library(ggplot2)
library(dplyr)

attach(Data)

x=Age
y=Taille

summary(Data)

# Histogramme univarié de la Taille
hist(Data$Taille,
      main = "Histogramme de la taille",
      xlab = "Taille (en cm)", ylab = "Nombre de personne",
      col = "lightblue",
      border = "white")

# 1) Analyse descriptive
summary(x)      # Statistiques descriptives sur l'âge
sd(x)           # Écart-type de l'âge
var(x)          # Variance de l'âge

summary(y)      # Statistiques descriptives sur la taille
sd(y)           # Écart-type de la taille
var(y)          # Variance de la taille

# 2) Graphe de dispersion
plot(x, y,
      main = "Taille en fonction de l'âge",
      xlab = "Âge (en années)",
      ylab = "Taille (en cm)",
      col = "blue",
      pch = 19)

# 3) Ajout d'une droite de régression linéaire simple (optionnelle pour visualiser la tendance)
model <- lm(y ~ x)
abline(model, col = "red", lwd = 2)

# 4) Afficher la courbe de tendance polynomiale (degré 2 par exemple)
model_poly <- lm(y ~ poly(x, 2))
curve(predict(model_poly, newdata = data.frame(x = x)),
      add = TRUE, col = "green", lwd = 2)
legend("topleft", legend = c("Données", "Régression linéaire", "Régression polynomiale (ordre 2)"),
      col = c("blue", "red", "green"), lwd = 2, pch = c(19, NA, NA), lty = c(NA, 1, 1))
```

```

# Boite a moustache comparent la taille en fonction du pays
ggplot(data = Data, aes(x = Pays, y = Taille, fill = Pays)) +
  geom_boxplot() +
  scale_fill_manual(values = c("green", "yellow")) +
  labs(
    title = "Distribution de la taille par pays",
    x = "Pays",
    y = "Taille")

# Boîte à moustache comparent la taille en fonction du sexe et du Pays
ggplot(Data, aes(x = Pays, y = Taille, fill = Sexe)) +
  geom_boxplot() +
  labs(
    title = "Boîte à moustache de la taille en fonction du sexe et du Pays",
    x = "Pays",
    y = "Taille")

# Séparation des données par sexe
Data_boys <- subset(Data, Sexe == "Boys")
Data_girls <- subset(Data, Sexe == "Girls")

# Graphique pour les garçons
plot(Data_boys$Age, Data_boys$Taille,
     main = "Taille en fonction de l'âge - Garçons",
     xlab = "Âge (années)",
     ylab = "Taille (cm)",
     col = "blue",
     pch = 19)

# Régression linéaire garçons
model_boys <- lm(Taille ~ Age, data = Data_boys)
abline(model_boys, col = "green", lwd = 2)

# Régression polynomiale de degrés 2 (garçons)
model_poly_boys <- lm(Taille ~ poly(Age, 2), data = Data_boys)
curve(predict(model_poly_boys, newdata = data.frame(Age = x)),
      from = min(Data_boys$Age), to = max(Data_boys$Age),
      add = TRUE, col = "red", lwd = 2)

# Régression polynomiale de degrés 3 (garçons)
model_poly3_boys <- lm(Taille ~ poly(Age, 3), data = Data_boys)
curve(predict(model_poly3_boys, newdata = data.frame(Age = x)),
      from = min(Data_boys$Age), to = max(Data_boys$Age),
      add = TRUE, col = "pink", lwd = 2)

legend("topleft",
      legend = c("Points", "Régression linéaire", "Régression polynomiale (ordre 2)", "Régression polynomiale (ordre 3)"),
      col = c("blue", "green", "red", "pink"),
      pch = c(19, NA, NA, NA), lty = c(NA, 1, 1, 1), lwd = 2)
# ----- 2) Graphique pour les filles -----
plot(Data_girls$Age, Data_girls$Taille,
     main = "Taille en fonction de l'âge - Filles",
     xlab = "Âge (années)",
     ylab = "Taille (cm)",
     col = "deeppink",
     pch = 19)

```

```

# Régression linéaire filles
model_girls <- lm(Taille ~ Age, data = Data_girls)
abline(model_girls, col = "green", lwd = 2)

# Régression polynomiale de degrés 2 (filles)
model_poly_girls <- lm(Taille ~ poly(Age, 2), data = Data_girls)
curve(predict(model_poly_girls, newdata = data.frame(Age = x)),
      from = min(Data_girls$Age), to = max(Data_girls$Age),
      add = TRUE, col = "purple", lwd = 2)

# Régression polynomiale de degrés 3 (filles)
model_poly3_girls <- lm(Taille ~ poly(Age, 3), data = Data_girls)
curve(predict(model_poly3_girls, newdata = data.frame(Age = x)),
      from = min(Data_girls$Age), to = max(Data_girls$Age),
      add = TRUE, col = "gold", lwd = 2)

legend("topleft",
      legend = c("Points", "Régression linéaire", "Régression polynomiale (ordre 2)", "Régression polynomiale (ordre 3)"),
      col = c("deeppink", "green", "purple", "gold"),
      pch = c(19, NA, NA, NA), lty = c(NA, 1, 1, 1), lwd = 2)

```

Essai 2

2. Filtrer les données pour chaque groupe

```

Data_pf_boys <- subset(Data, Pays == "French Polynesia" & Sexe == "Boys")
Data_pf_girls <- subset(Data, Pays == "French Polynesia" & Sexe == "Girls")
Data_sl_boys <- subset(Data, Pays == "Saint Lucia" & Sexe == "Boys")
Data_sl_girls <- subset(Data, Pays == "Saint Lucia" & Sexe == "Girls")

```

3. Modèles polynomiaux de degrés 2

```

model_pf_boys <- lm(Taille ~ poly(Age, 2), data = Data_pf_boys)
model_pf_girls <- lm(Taille ~ poly(Age, 2), data = Data_pf_girls)
model_sl_boys <- lm(Taille ~ poly(Age, 2), data = Data_sl_boys)
model_sl_girls <- lm(Taille ~ poly(Age, 2), data = Data_sl_girls)

```

4. Tracé des courbes

Créer un vecteur d'âges pour la prédiction

```
age_seq <- seq(5, 19, by = 0.1)
```

Générer les prédictions pour chaque groupe

```

pred_pf_boys <- data.frame(
  Age = age_seq,
  Taille = predict(model_pf_boys, newdata = data.frame(Age = age_seq)),
  Groupe = "Polynésie française - Garçons"
)

pred_pf_girls <- data.frame(
  Age = age_seq,
  Taille = predict(model_pf_girls, newdata = data.frame(Age = age_seq)),
  Groupe = "Polynésie française - Filles"
)

pred_sl_boys <- data.frame(
  Age = age_seq,
  Taille = predict(model_sl_boys, newdata = data.frame(Age = age_seq)),
  Groupe = "Sainte-Lucie - Garçons"
)

pred_sl_girls <- data.frame(
  Age = age_seq,
  Taille = predict(model_sl_girls, newdata = data.frame(Age = age_seq)),
  Groupe = "Sainte-Lucie - Filles"
)

```

```

# Fusionner toutes les prédictions
predictions <- rbind(pred_pf_boys, pred_pf_girls, pred_sl_boys, pred_sl_girls)

# 5. Graphique ggplot avec couleurs personnalisées
ggplot(predictions, aes(x = Age, y = Taille, color = Groupe)) +
  geom_line(size = 1.2) +
  scale_color_manual(
    values = c(
      "Polynésie française - Garçons" = "blue",
      "Polynésie française - Filles" = "pink",
      "Sainte-Lucie - Garçons" = "lightblue",
      "Sainte-Lucie - Filles" = "violet"
    )
  ) +
  labs(
    title = "Courbes de croissance polynomiales de degré 2",
    x = "Âge (en années)",
    y = "Taille (cm)"
  ) +
  theme_minimal() +
  theme(legend.title = element_blank())

# Courbes de croissances polynomiales de degres 3 a comparer avec celle de degres 2

# 2. Filtrer les données pour chaque groupe
Data_pf_boys <- subset(Data, Pays == "French Polynesia" & Sexe == "Boys")
Data_pf_girls <- subset(Data, Pays == "French Polynesia" & Sexe == "Girls")
Data_sl_boys <- subset(Data, Pays == "Saint Lucia" & Sexe == "Boys")
Data_sl_girls <- subset(Data, Pays == "Saint Lucia" & Sexe == "Girls")

# 3. Modèles polynomiaux de degré 3
model_pf_boys_3 <- lm(Taille ~ poly(Age, 3), data = Data_pf_boys)
model_pf_girls_3 <- lm(Taille ~ poly(Age, 3), data = Data_pf_girls)
model_sl_boys_3 <- lm(Taille ~ poly(Age, 3), data = Data_sl_boys)
model_sl_girls_3 <- lm(Taille ~ poly(Age, 3), data = Data_sl_girls)

# 4. Créer un vecteur d'âges pour la prédiction
age_seq <- seq(5, 19, by = 0.1)

# Générer les prédictions uniquement pour le degré 3
pred_pf_boys <- data.frame(
  Age = age_seq,
  Taille = predict(model_pf_boys_3, newdata = data.frame(Age = age_seq)),
  Groupe = "Polynésie française - Garçons"
)

pred_pf_girls <- data.frame(
  Age = age_seq,
  Taille = predict(model_pf_girls_3, newdata = data.frame(Age = age_seq)),
  Groupe = "Polynésie française - Filles"
)

pred_sl_boys <- data.frame(
  Age = age_seq,
  Taille = predict(model_sl_boys_3, newdata = data.frame(Age = age_seq)),
  Groupe = "Sainte-Lucie - Garçons"
)

pred_sl_girls <- data.frame(
  Age = age_seq,
  Taille = predict(model_sl_girls_3, newdata = data.frame(Age = age_seq)),
  Groupe = "Sainte-Lucie - Filles"
)

# Fusionner toutes les prédictions
predictions <- rbind(pred_pf_boys, pred_pf_girls, pred_sl_boys, pred_sl_girls)

```

```

# 5. Graphique ggplot avec couleurs personnalisées uniquement pour degré 3
ggplot(predictions, aes(x = Age, y = Taille, color = Groupe)) +
  geom_line(size = 1.2) +
  scale_color_manual(
    values = c(
      "Polynésie française - Garçons" = "blue",
      "Polynésie française - Filles" = "pink",
      "Sainte-Lucie - Garçons" = "lightblue",
      "Sainte-Lucie - Filles" = "violet"
    )
  ) +
  labs(
    title = "Courbes de croissance polynomiales de degré 3",
    x = "Âge (en années)",
    y = "Taille (cm)"
  ) +
  theme_minimal() +
  theme(legend.title = element_blank())

#3 tranche d'âges "essaie" ça marche
# Découpe l'âge en tranches de 1 an
Data <- Data %>%
  mutate(tranche_age = cut(Age, breaks = seq(5, 19, by = 1), include.lowest = TRUE))

# Calcule les quantiles de taille par groupe et tranche d'âge
croissance <- Data %>%
  group_by(Pays, Sexe, tranche_age) %>%
  summarise(
    Q1 = quantile(Taille, 0.25, na.rm = TRUE),
    Mediane = quantile(Taille, 0.50, na.rm = TRUE),
    Q3 = quantile(Taille, 0.75, na.rm = TRUE),
    .groups = "drop"
  ) %>%

# Calcul âge central = moyenne des bornes de la tranche
mutate(
  Age_central = as.numeric(sub("\\((.+),(.+)\\)", "\\1", tranche_age)) + 0.5,
  Groupe = paste(Sexe, Pays, sep = "_")
)

# Supprime les lignes avec NA
croissance_net <- croissance %>%
  filter(!is.na(Q1), !is.na(Mediane), !is.na(Q3), !is.na(Age_central))

# Graphique avec couleurs personnalisées pour Q1, Médiane, Q3
ggplot(croissance_net, aes(x = Age_central)) +
  geom_line(aes(y = Mediane, color = "Médiane", group = Groupe), linewidth = 0.5) +
  geom_line(aes(y = Q1, color = "1er Quartile", group = Groupe), linetype = "dashed", linewidth = 0.5) +
  geom_line(aes(y = Q3, color = "3e Quartile", group = Groupe), linetype = "dashed", linewidth = 0.5) +
  facet_grid(Sexe ~ Pays) +
  scale_color_manual(
    values = c(
      "1er Quartile" = "violet",
      "Médiane" = "red",
      "3e Quartile" = "lightblue"
    )
  ) +
  labs(
    title = "Courbes de croissance (Médiane, Q1, Q3)",
    x = "Âge (années)",
    y = "Taille (cm)",
    color = "Courbe"
  ) +
  theme_minimal()

```

#Superposition des 4 courbes médianes

```
# 1) Ajouter tranche d'âge (de 5 à 19 ans par tranche de 1 an)
Data <- Data %>%
mutate(tranche_age = cut(Age, breaks = seq(5, 19, by = 1), include.lowest = TRUE))

# 2) Calculer quantiles par Pays, Sexe, tranche_age
croissance <- Data %>%
  group_by(Pays, Sexe, tranche_age) %>%
  summarise(
    Mediane = quantile(Taille, 0.5, na.rm = TRUE),
    .groups = "drop"
  ) %>%
  # Extraire l'âge central de la tranche
  mutate(
    Age_central = as.numeric(sub("\\((.+),(.+)\\]", "\\1", tranche_age)) + 0.5,
    Groupe = paste(Sexe, Pays, sep = "_")
  ) %>%
  filter(!is.na(Mediane), !is.na(Age_central))

# 3) Choisir 4 groupes à tracer (modifier selon tes groupes)
groupes_a_tracer <- unique(croissance$Groupe)[1:4] # Prend les 4 premiers groupes disponibles

croissance_filtre <- croissance %>% filter(Groupe %in% groupes_a_tracer)

# 4) Tracer les 4 médianes sur un même graphique
ggplot(croissance_filtre, aes(x = Age_central, y = Mediane, color = Groupe)) +
  geom_line(size = 1) +
  labs(
    title = "Superposition des 4 courbes médianes",
    x = "Âge (années)",
    y = "Taille médiane (cm)",
    color = "Groupe"
  ) +
  theme_minimal()
```