

# SAE

## Régression sur données réelles



Université de Paris

# Sommaire

## **1** - Comparaison des tailles entre garçons

**1.1** - Analyse descriptive chez les garçons

**1.2** - Modèles linéaires

**1.3** - Analyse descriptive

## **2** - Comparaison des tailles entre garçons

**2.1** - Analyse descriptive

**2.2** - Analyse descriptive

**2.3** - Analyse descriptive

# Introduction

En 2020, une étude internationale a été réalisée dans plus de 90 pays afin d'obtenir des informations sur les courbes de croissance des filles et des garçons âgés de 5 à 19 ans. La méthodologie de cette étude est la suivante : pour chaque pays, 200 médecins ont été sélectionnés de manière aléatoire sur l'ensemble du territoire. Chaque médecin a ensuite choisi de manière aléatoire 20 patients âgés de 5 à 19 ans, et a enregistré leur âge et leur taille. L'âge a été enregistré par tous les praticiens en arrondissant à 1 chiffre après la virgule (en années décimales), et la taille en arrondissant à 1 chiffre après la virgule (en cm). Pour chaque pays inclus dans l'étude, les données concernant l'âge et la taille de 4000 enfants (2000 de chaque sexe) étaient disponibles.

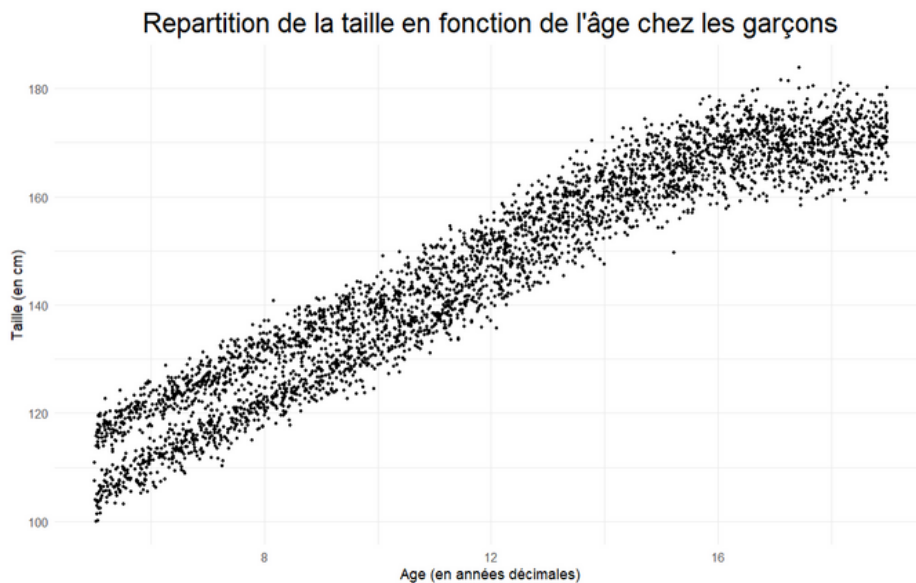
Dans ce rapport nous nous baserons sur 2 pays avec un échantillon de 8000 enfants issus :

- Des Fidji (Fiji) : pays du Pacifique Sud, composé de plus de 300 îles.
- L'Eswatini (Swaziland) : petite monarchie situé au sud de l'Afrique

Nous étudierons si il existe une différence significative ou non des tailles selon le pays d'origines des individus.

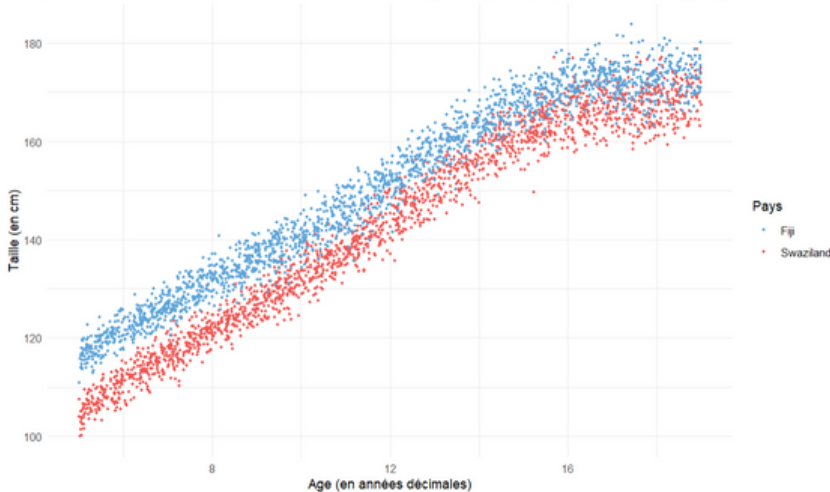
# Observation chez les garçons

## 1.1) Analyse descriptive chez les garçons



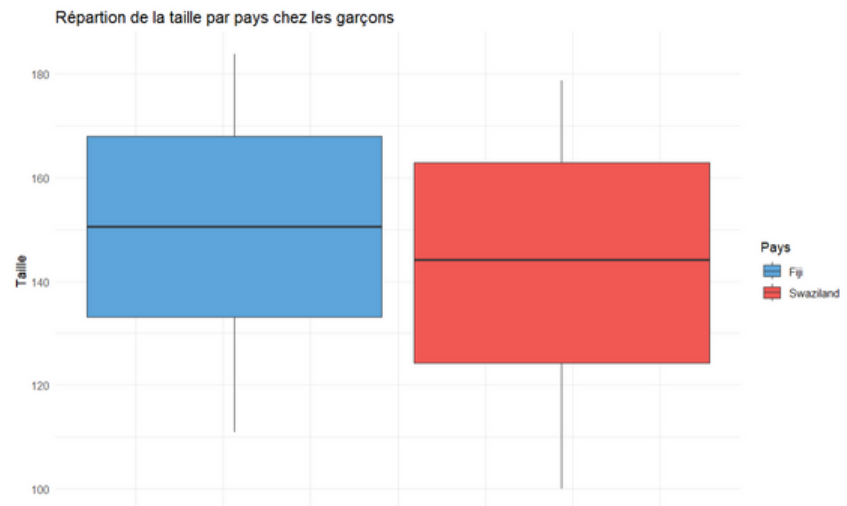
Ce nuage de points représentant la distribution des tailles mesurées selon l'âge chez les garçons des différents pays. On remarque un écart flagrant chez les garçons entre 5 et 10 ans.

Repartition de la taille en fonction de l'âge et par pays chez les garçons

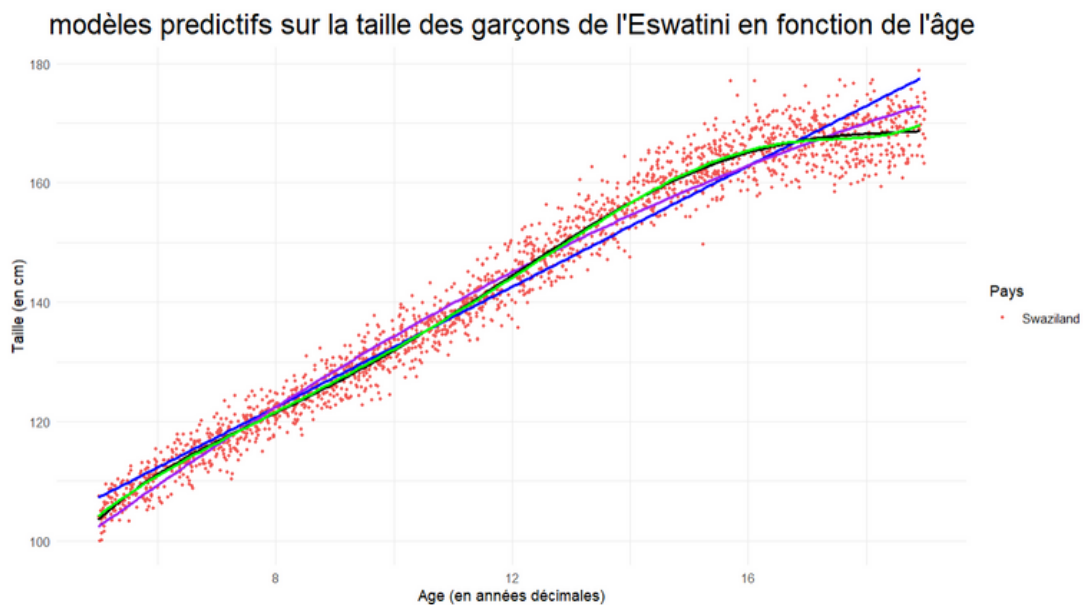


D'après le même nuage de points, nous pouvons voir des écarts de tailles entre les garçons de l'observation qui peuvent être expliqués en fonction du pays. En effet, on remarque que les garçons des Fidji représentés par des points bleus sont en général plus grands que ceux de l'Eswatini représentés en points rouge.

Grâce aux boîtes à moustache, nous pouvons voir que la taille médian des garçons des Fidji est légèrement plus élevée que celle de l'Eswatini, 1m50 contre 1m45. Visuellement, nous pouvons également noter une dispersion de tailles plus grande en Eswatini.



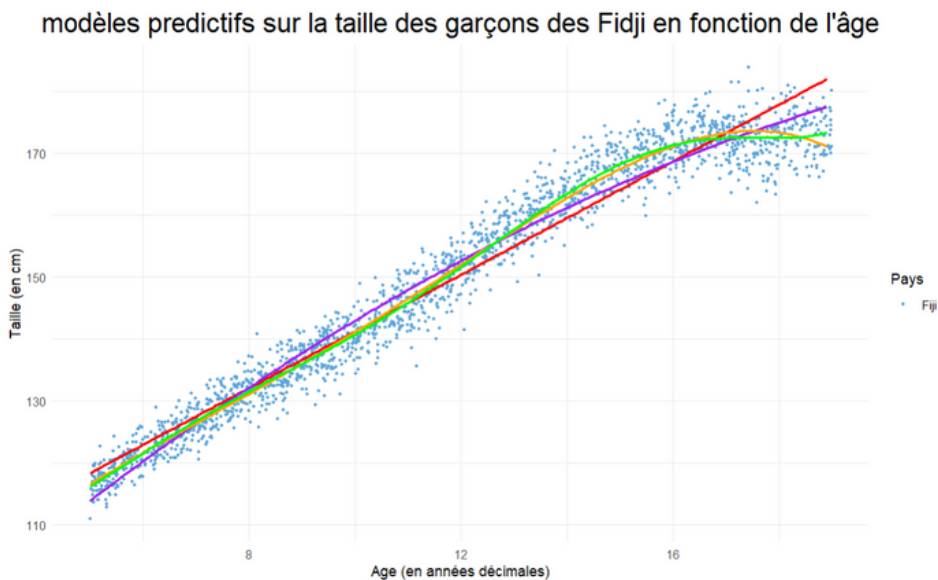
## 1.2) Modèles linéaires



Sur ce graphique nous avons 4 courbes qui représentent 4 ajustements polynomiaux de degré différents :

- En bleu : degré 1
- En violet : degré 2
- En noir : degré 5
- En vert : degré 7

Le modèle le plus pertinent est illustré par la courbe verte qui prédit le mieux la tailles des garçons swazis d'après le critère d'information d'Akaike (AIC). Cette courbe pourrait nous permettre de prédire et de comprendre l'évolution de tailles des swazis en fonction de l'âge.



Sur ce graphique nous avons 4 courbes qui représentent 4 ajustements polynomiaux de degré différents :

- En rouge : degré 1
- En violet : degré 2
- En orange : degré 4
- En vert : degré 6

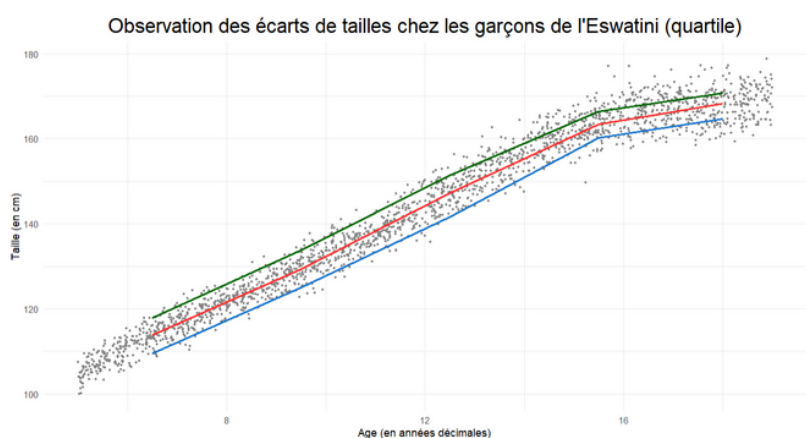
Le modèle le plus pertinent est illustré par la courbe verte qui prédit le mieux la tailles des garçons fidjiens d'après le critère d'information d'Akaike (AIC). Notre échantillon se limite à une tranche d'âge compris entre 5 et 19 ans, mais nous pourrons grâce à cette courbe prédire la taille d'un homme fidjiens âgé de plus de 19 ans.

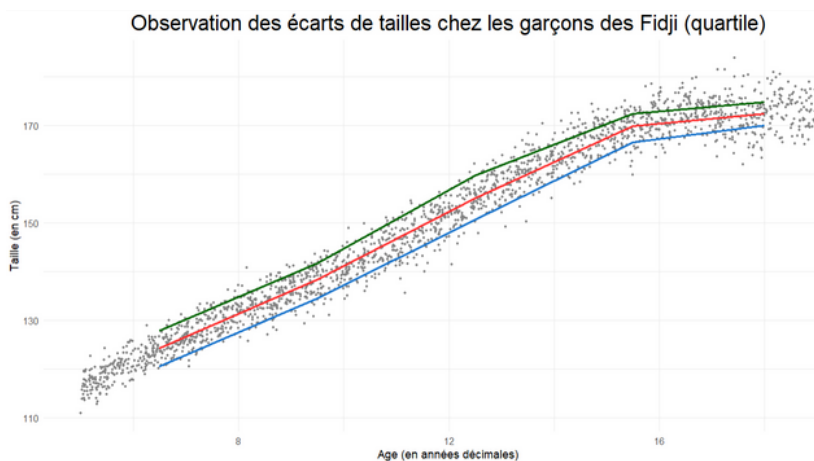
## 1.3) Observation des tailles des garçons selon les quartiles

Sur ce graphique est tracé 3 courbes selon 5 classes d'âge [5;8], [8;11], [11;14], [14;17] et [17;19] :

- La bleue : le 1er quartile
- La rouge : la médiane
- La verte : le 3ème quartile

Cette illustration peut nous donner un aperçu de la dispersions des tailles selon les classes d'âges. Mais aussi une observation d'une taille à un âge donné : pour 12 ans la taille médian est 1m45.

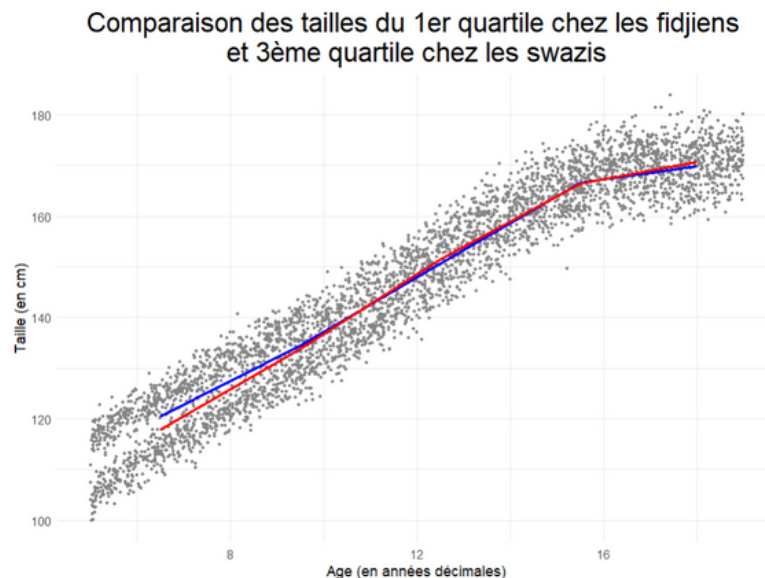




Sur ce graphique est tracé 3 courbes selon 5 classes d'âge énoncé précédemment :

- La bleue : le 1er quartile
- La rouge : la médiane
- La verte : le 3ème quartile

Graphiquement, nous pouvons voir qu'à 16 ans 1/4 des garçons font environ 1m66 ou moins.



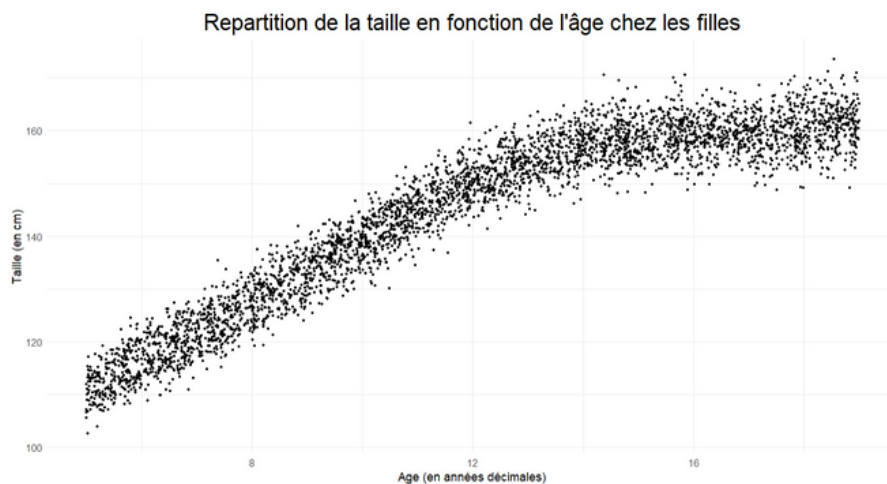
Sur ce graphique nous avons 2 courbes :

- Celle en bleue représente le 1er quartile des tailles des fidjiens, 25% des tailles sont inférieures ou égales au 1er quartile.
- Celle en rouge représente le 3er quartile des tailles des swazis, 75% des tailles sont inférieurs ou égales au 3ème quartile.

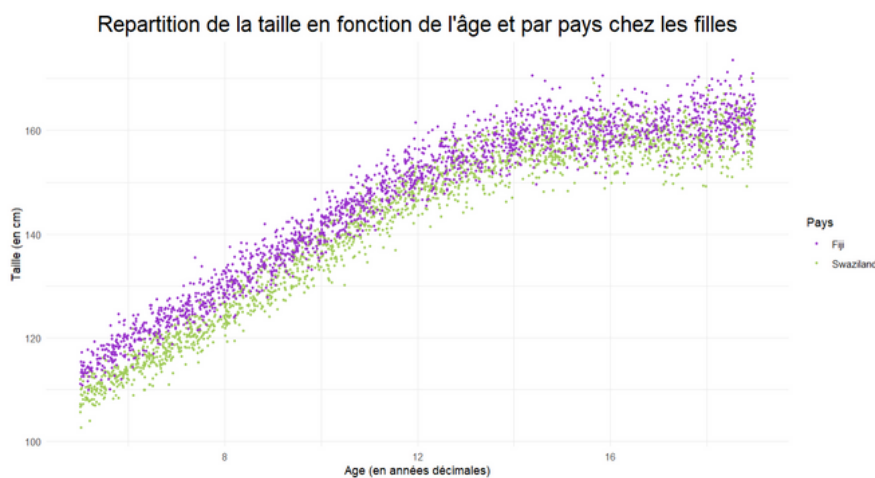
On peut admettre d'après cette analyse que 75% des garçons swazis sont aussi grands que 25% des garçons fidjiens. Montrant un grand écart de taille entre les 2 pays en faveur des Fidji.

# Observation chez les filles

## 1.1) Analyse descriptive chez les filles



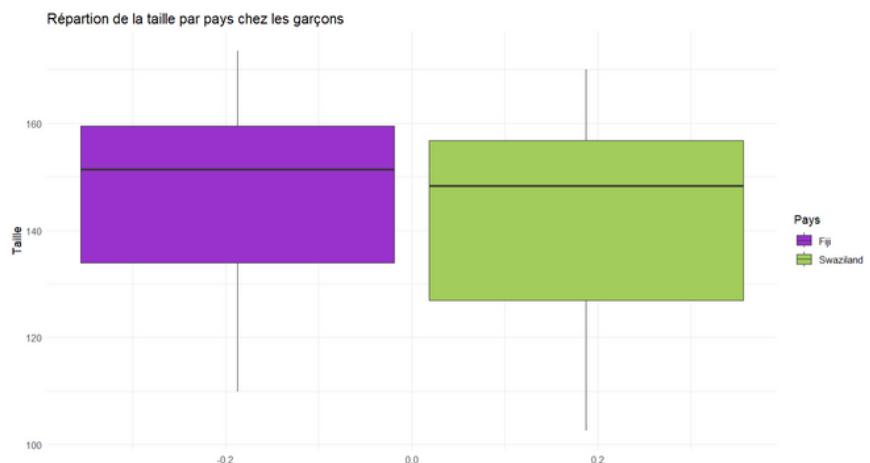
Ce nuage de points représentant la distribution des tailles mesurées selon l'âge chez les filles des 2 différents pays. Contrairement aux garçons les observations sont moins dispersées, on peut supposer que les filles des 2 pays ont des tailles similaires.



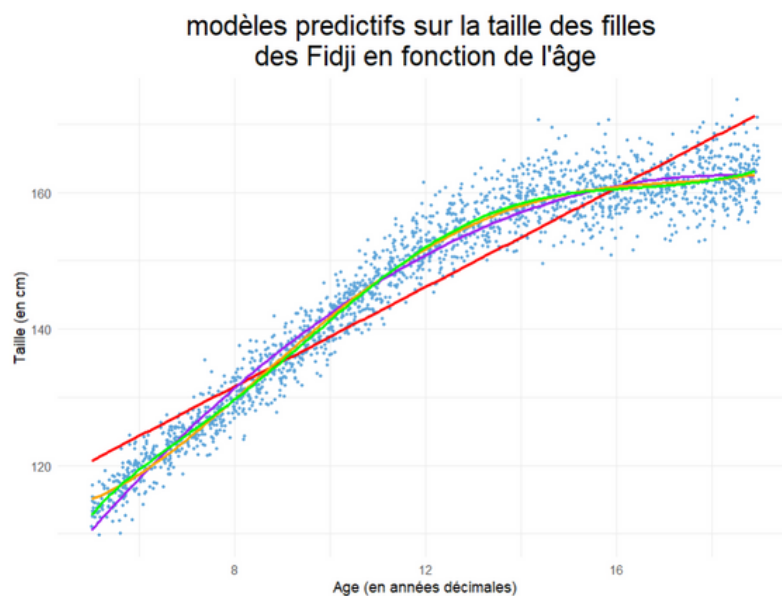
En distinguant les points du nuage selon les pays, on remarque la même tendance que chez les garçons. Les filles issues des Fidji sont généralement plus grandes que les filles de l'Eswatini, mais les différences de tailles sont moins flagrantes.



En distinguant les points du nuage selon les pays, on remarque la même tendance que chez les garçons. Les filles issues des Fidji sont généralement plus grandes que les filles de l'Eswatini, mais les différences de tailles sont moins flagrantes.



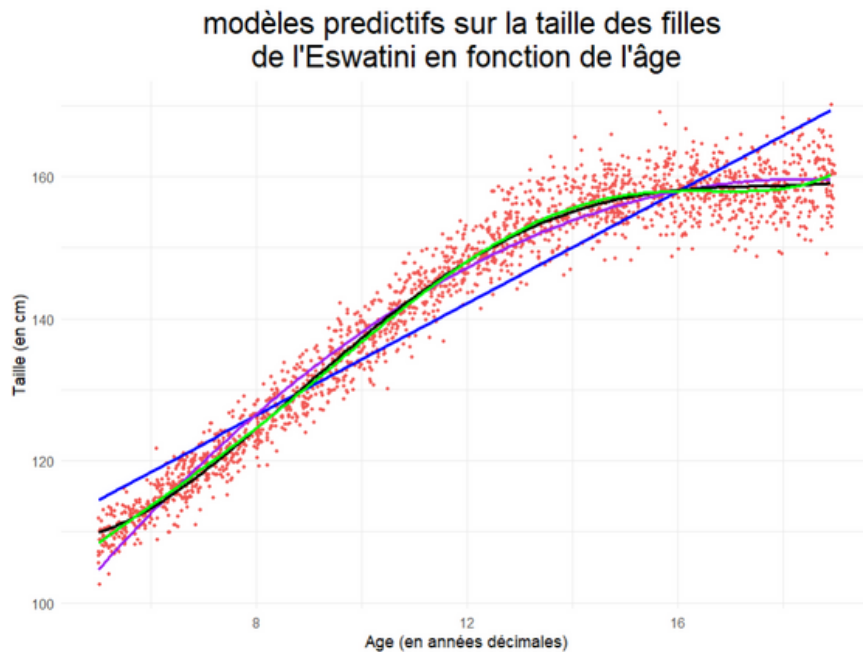
## 1.2) Modèles linéaires



Sur ce graphique nous avons 4 courbes qui représentent 4 ajustements polynomiaux de degré différents :

- En rouge : degré 1
- En violet : degré 2
- En orange : degré 4
- En vert : degré 6

Le modèle le plus pertinent est illustré par la courbe verte qui prédit le mieux la tailles des filles fidjiennes d'après le critère d'information d'Akaike (AIC). Connaître ce modèle nous permettrait de prédire la taille d'une femme fidjiens pour un âge donné.

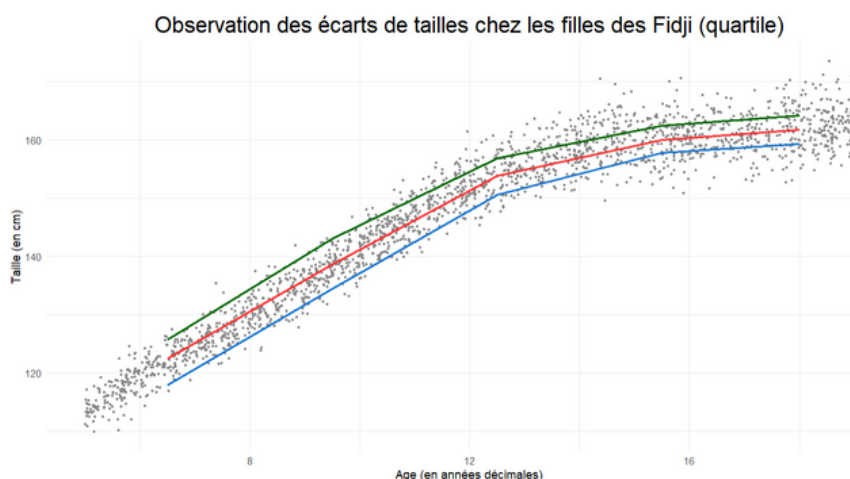


Sur ce graphique nous avons 4 courbes qui représentent 4 ajustements polynomiaux de degré différents :

- En bleu : degré 1
- En violet : degré 2
- En noir : degré 4
- En vert : degré 5

Le modèle le plus pertinent est illustré par la courbe verte qui prédit le mieux la tailles des filles swazis d'après le critère d'information d'Akaike (AIC). Cette courbe pourrait nous permettre de prédire et de comprendre l'évolution de tailles des swazies en fonction de l'âge.

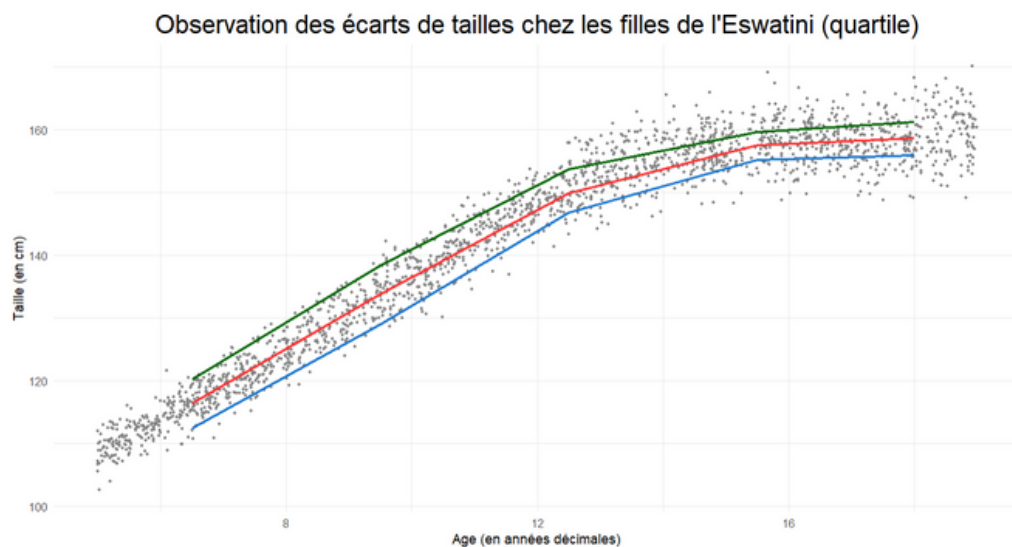
## 2.3) Observation des filles selon le 1er, 2ème et 3ème quartile



Comme pour les garçons ce graphique s'est construit autour de 5 classes d'âges [5;8], [8;11], [11;14], [14;17] et [17;19], et on y voit 3 courbes :

- La bleu : le 1er quartile
- La rouge : la médiane
- La verte : le 3ème quartile

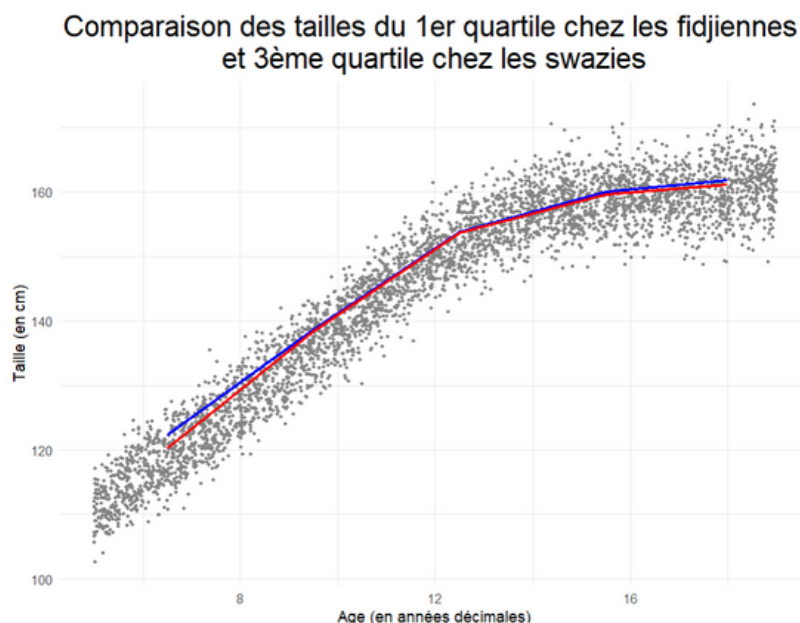
On peut voir par exemple qu'à 12 ans la taille médiane des filles des Fidji était d'environ 1m50.



Comme pour les garçons ce graphique s'est construit autour de 5 classes d'âges [5;8], ]8;11], ]11;14], ]14;17] et ]17;19], et on y voit 3 courbes :

- La bleu : le 1er quartile
- La rouge : la médiane
- La verte : le 3ème quartile

On peut voir par exemple qu'à 12 ans la taille médiane des filles des Fidji était d'environ 1m50.



Sur ce graphique nous avons 2 courbes :

- Celle en bleue représente la taille médian des fidjiennes en fonction de l'âge, 50% des tailles sont inférieures/supérieur à la médiane.
- Celle en rouge représente le 3er quartile des tailles des swazis, 75% des tailles sont inférieurs ou égales au 3ème quartile.

Nous pouvons dire que la moitié des filles fidjiennes sont aussi grande voire plus encore que 75% des jeunes swazies. Montrant encore une fois un écart de taille important de tailles.

# Conclusion

Dans le cadre d'une étude sur la taille d'individus issus de 2 pays : Fidji, Eswatini, nous voulions voir l'évolution de la taille en fonction de l'âge selon le pays mais aussi constater s'il existait des différences de tailles flagrantes entre ces populations. Notre étude s'est basé sur un échantillon de 8000 observations, réparties en fonction des pays : 4000 fidjiens/4000 swazis parmi lesquels nous avons 2000 fidjiens/2000 fidjiennes et 2000 swazis/2000 swazies.

Notre étude montre bel et bien une différence de taille selon si nous étions fidjien ou swazi. En effet, d'après les graphiques qui ont été réalisé nous avons pu observer que les garçons et les filles des Fidji étaient généralement plus grands en taille que leur homologue swazi. Ces écarts de tailles étaient plus visibles chez les garçons, car nous avons estimé que  $\frac{3}{4}$  des garçons swazis étaient aussi grands qu' $\frac{1}{4}$  des garçons des Fidji. Et chez les filles  $\frac{1}{2}$  des fidjiennes étaient au moins aussi grandes que  $\frac{3}{4}$  des swazies.

As part of a study on the height of individuals from two countries, Fiji and Swaziland, we wanted to examine the relationship between height and age within each country and also determine if there were significant differences in height between these populations. Our study was based on a sample of 8000 observations, divided by country: 4000 fijians and 4000 swazis, including 2000 fijian boys, 2000 fijian girls, 2000 swazi boys, and 2000 swazi girls. Our study indeed demonstrates a difference in height between fijians and swazis. According to our observations, we observed that both Fijian boys and girls were generally taller than their swazi counterparts. These height disparities were more pronounced among boys, as we estimated that  $\frac{3}{4}$  of swazi boys were as tall as only  $\frac{1}{4}$  of fijian boys. Among girls, at least half of fijian girls were taller than  $\frac{3}{4}$  of swazi girls.