

IUT Grand Ouest Normandie

Bachelor Universitaire de Technologie
Science des Données
Campus de Lisieux

Science des Données 1 - SAE 2.03
Régression sur données réelles

Thématique

Étude des manchots des archipels Palmer en Antarctique



Auteurs

BOUCHER Nathan
BRAUD Ewen
GAMONDELE Maxime

Table des matières

Introduction	3
1 Étude du poids des manchots	4
1.1 Le poids des manchots en fonction de la longueur des nageoires	4
1.2 Le poids des manchots en fonction de la longueur de la crête supérieur du bec	4
1.3 Le poids des manchots en fonction de la profondeur de la crête supérieur du bec	5
2 Étude du poids des manchots, selon le sexe	6
2.1 Le poids des manchots en fonction de la longueur des nageoires, selon le sexe	6
2.2 Le poids des manchots en fonction de la longueur de la crête supérieur du bec, selon le sexe	6
2.3 Le poids des manchots en fonction de la profondeur de la crête, selon le sexe	7
3 Étude du poids des manchots, selon l'espèce	8
3.1 Le poids des manchots en fonction de la longueur des nageoires, selon l'espèce	8
3.2 Le poids des manchots en fonction de la longueur de la crête supérieur du bec, selon l'espèce	8
3.3 Le poids des manchots en fonction de la profondeur de la crête supérieur du bec, selon l'espèce	9
Conclusion	10
4 Annexe	11
Table des figures	12

Introduction

Un ornithologue a récupéré un jeu de données portant sur un échantillon de 244 manchots provenant des archipels Palmer en Antarctique. Voici les variables contenues dans le fichier (cf. Fig. 1) :

- L'espèce (**Species** - Adélie, Chinstrap ou Gentoo);
- L'île d'origine (**Island** - Dream, Torgersen ou Biscoe);
- La longueur de la crête supérieure du bec (**bill_length_mm**);
- La profondeur de la crête supérieure du bec (**bill_depth_mm**);
- La longueur des nageoires (**flipper_length_mm**);
- Le poids (**body_mass_g**);
- Le sexe (**sex**);
- L'année d'étude (**year**).

Fig. 1. Variables du jeu de données

Nous proposons d'effectuer une analyse exploratoire sur le poids des manchots en fonction de différents paramètres. Notre analyse sera divisée en trois parties distinctes. La première partie portera sur l'étude du poids des manchots en fonction de trois variables : la longueur de leurs nageoires, la longueur de la crête supérieure du bec et la profondeur de la crête supérieure du bec. La seconde partie se concentrera sur les mêmes variables explicatives, mais cette fois en fonction du sexe des individus. Enfin, la dernière partie selon l'espèce des manchots.

1 Étude du poids des manchots

Cette première partie abordera l'analyse des données concernant le poids des manchots en fonction des variables suivantes : la longueur de leurs nageoires, la longueur de la crête supérieure du bec et la profondeur de la crête.

1.1 Le poids des manchots en fonction de la longueur des nageoires

Sur ce premier graphique (cf. Fig. 2), nous observons un coefficient de corrélation linéaire de Pearson de 0,87, indiquant une très forte association entre les deux variables. Cette relation est visible à travers les points qui s'alignent étroitement avec la droite des moindres carrés en rouge. Le coefficient directeur de la droite nous renseigne que pour chaque augmentation de 1 cm de la longueur des nageoires, le poids augmente en moyenne de 0,497 kg. De plus, le coefficient de détermination nous indique que la variable de la longueur des nageoires explique 75,9 % de la variabilité du poids.

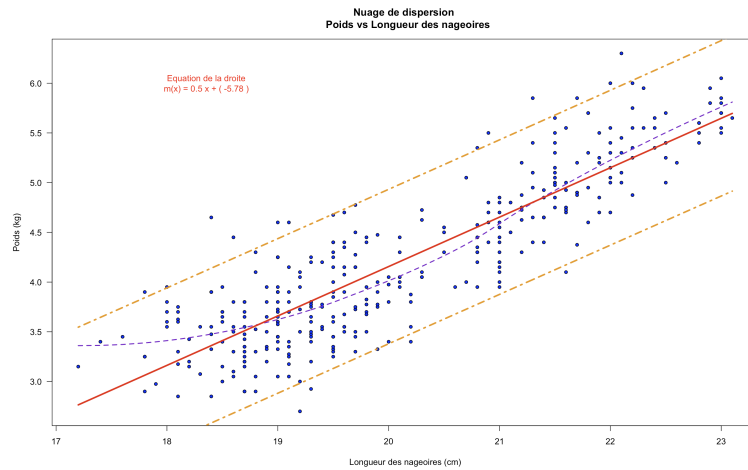


Fig. 2. Le poids des manchots en fonction de la longueur des nageoires

1.2 Le poids des manchots en fonction de la longueur de la crête supérieure du bec

Ce graphique (cf. Fig. 3), silaire au précédent, vise à expliquer le poids en fonction de la longueur de la crête supérieure du bec. Nous observons un coefficient de corrélation linéaire de Pearson de 0.59, ce qui indique une association linéaire modérée entre ces variables. L'augmentation de 1 cm de la longueur de la crête entraîne une augmentation moyenne de 0.874 kg. De plus, la variable longueur de la crête explique 35.42 % de la variabilité de la variable poids.

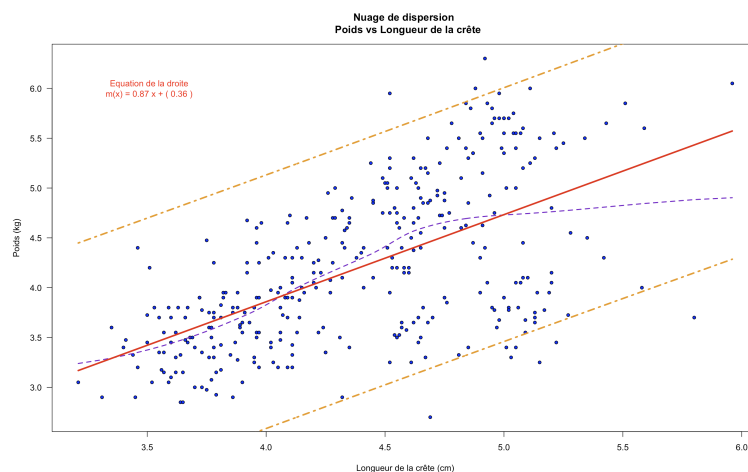


Fig. 3. Le poids des manchots en fonction de la longueur de la crête supérieur du bec

1.3 Le poids des manchots en fonction de la profondeur de la crête supérieur du bec

Ci-dessous (cf. Fig. 4), le nuage de points présente une allure différente des précédents. Nous observons une distinction entre deux zones, ce qui suggère que la variable sexe pourrait affecter les résultats. Le coefficient de corrélation linéaire de Pearson est de -0.47, indiquant une association modérée entre les variables. L'augmentation de 1 cm de la profondeur de la crête induit une augmentation moyenne de -1.916 kg. Selon le coefficient de détermination, la variable de profondeur de la crête explique 22,27 % de la variabilité du poids non prise en compte dans le modèle constant.

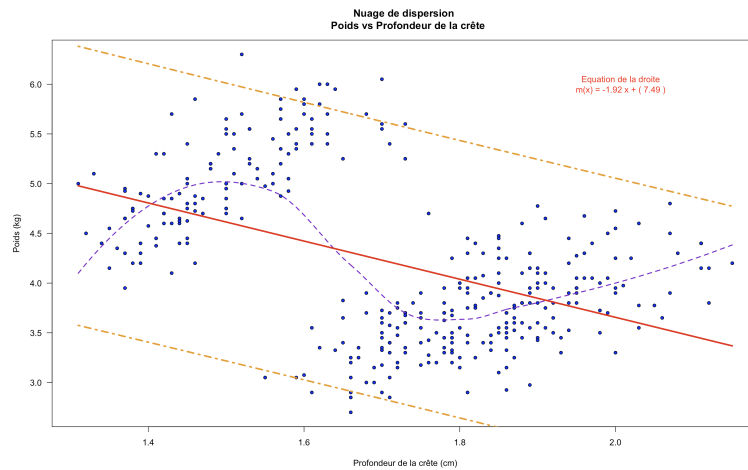


Fig. 4. Le poids des manchots en fonction de la profondeur de la crête supérieur du bec

2 Étude du poids des manchots, selon le sexe

Après avoir étudié le poids des manchots, nous allons maintenant effectuer une analyse similaire en tenant compte de la variable sexe.

2.1 Le poids des manchots en fonction de la longueur des nageoires, selon le sexe

Ce premier graphique(cf. Fig. 5), met en avant l'association linéaire entre les variables poids et longueur des nageoires selon le sexe. Chez les mâles, le coefficient de corrélation linéaire de Pearson est de 0.87 et chez les femelles de 0.88, indiquant ainsi une très forte association entre le poids et la longueur des nageoires des manchots. Pour les mâles, chaque augmentation de 1 cm de la longueur des nageoires entraîne en moyenne une augmentation de 0,469 kg, tandis que chez les femelles, cette augmentation moyenne est de 0,472 kg. Chez les mâles la variable Longueur des nageoires explique 74.91% de la variabilité de la variable Poids. Chez les femelles, un peu plus avec 78.29 %.

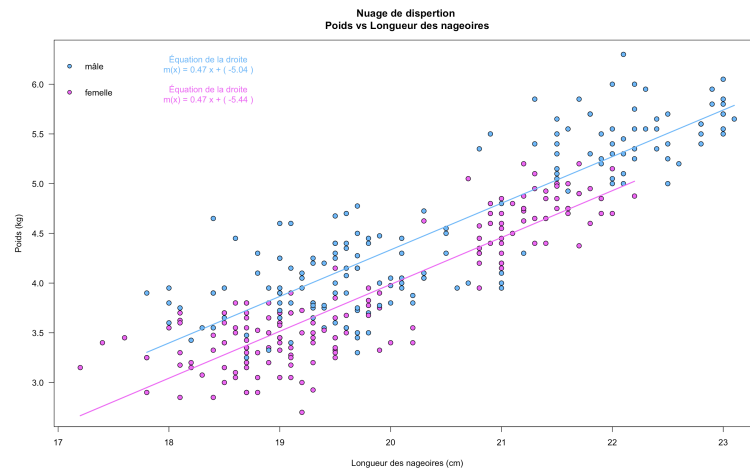


Fig. 5. Le poids des manchots en fonction de la longueur des nageoires, selon le sexe

2.2 Le poids des manchots en fonction de la longueur de la crête supérieur du bec, selon le sexe

Sur ce graphique(cf. Fig. 6), similaire au précédent, nous observons toujours deux entités mâles et femelles. Chez les mâles, le coefficient de corrélation linéaire de Pearson est de 0.48 et de 0.58 chez les femelles. Ce qui se traduit par une association modérée entre les variables poids et longueur de la crête, pour les deux sexes. De plus, une augmentation de 1 cm de la longueur de la crête entraîne une augmentation moyenne de 0.7 kg chez les manchots mâles, tandis que chez les femelles, cette augmentation moyenne est de 0.79 kg. Chez les mâles et les femelles, la longueur de la crête du bec explique respectivement 22.73 % et 33.79 % de la variabilité du poids.

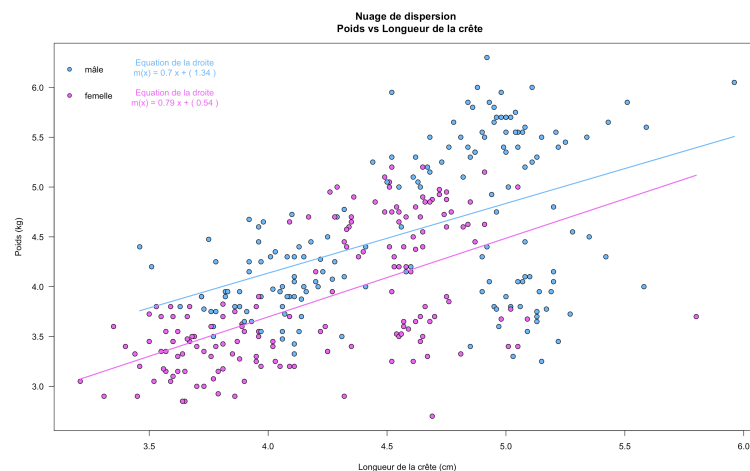


Fig. 6. Le poids des manchots en fonction de la longueur de la crête supérieur du bec, selon le sexe

2.3 Le poids des manchots en fonction de la profondeur de la crête, selon le sexe

Le nuage de dispersion ci-contre (cf. Fig. 7) contredit notre hypothèse émise dans la partie 1.3 (cf. section 1.3). Nous pensions trouver les mâles d'un côté et les femelles de l'autre. Nous observons un coefficient de Pearson de -0.76 chez les mâles et un de -0.75 chez les femelles. Soit une forte association entre les variables poids et profondeur de la crête. L'augmentation de 1 cm de la profondeur de la crête induit une diminution moyenne de -3.193 kg chez les manchots mâles et de -2.774 kg chez les manchots femelles. La profondeur de la crête explique respectivement 57,08 % et 55,92 % de la variabilité du poids chez les mâles et les femelles.

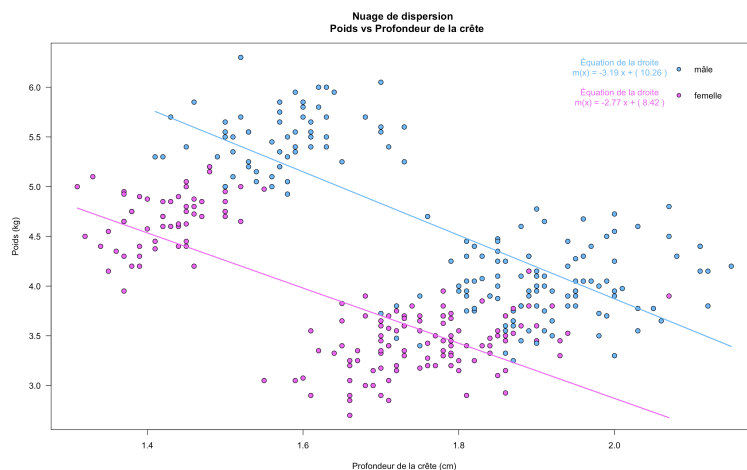


Fig. 7. Le poids des manchots en fonction de la profondeur de la crête, selon le sexe

3 Étude du poids des manchots, selon l'espèce

Après avoir examiné les variables explicatives sur plusieurs niveaux, nous allons maintenant nous concentrer sur l'étude du poids selon l'espèce des manchots.

3.1 Le poids des manchots en fonction de la longueur des nageoires, selon l'espèce

Ce premier nuage de points met en évidence les associations linéaires entre les variables poids et longueur de la crête selon les espèces (cf. Fig. 8). Pour l'espèce Adélie, Chinstrap et Gentoo, les coefficients de corrélation linéaire obtenus sont respectivement de 0.46, 0.64 et 0.71, ce qui indique une association modérée entre les variables poids et longueur des nageoires pour l'espèce Adélie, et une association forte pour les deux autres espèces. L'augmentation de 1 cm de Longueur des nageoires induit une augmentation moyenne de 0.327 kg chez l'espèce Adélie, une augmentation moyenne de 0.346 kg chez l'espèce Chinstrap et une augmentation moyenne de 0.542 kg chez l'espèce Gentoo. Chez l'espèce Adélie la variable Longueur des nageoires explique 21.61 % de la variabilité de la variable Poids, chez l'espèce Chinstrap 41.16 % et chez l'espèce Gentoo 50.6 %.

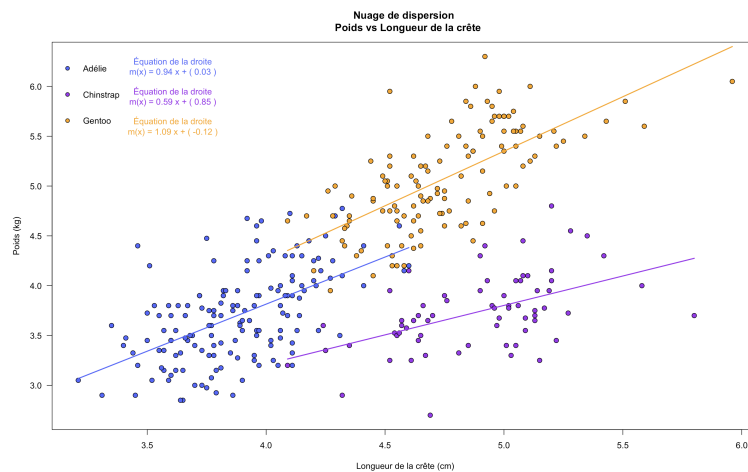


Fig. 8. Le poids des manchots en fonction de la longueur des nageoires, selon l'espèce

3.2 Le poids des manchots en fonction de la longueur de la crête supérieur du bec, selon l'espèce

Ce graphique (cf. Fig. 9) met en évidence l'association entre les variables poids et longueur de la crête selon l'espèce. On observe une association linéaire croissante pour chacune des espèces. Pour les espèces Adélie, Chinstrap et Gentoo, les coefficients de corrélation linéaires obtenus sont respectivement de 0,54, 0,51 et 0,67, indiquant une association linéaire modérée pour les deux premières espèces et forte pour les manchots Gentoo. Chaque augmentation de 1 cm de la longueur de la crête entraîne en moyenne une augmentation de 0,937 kg pour l'espèce Adélie, de 0,591 kg pour l'espèce Chinstrap et de 1,076 kg pour l'espèce Gentoo. De plus, la variable de la longueur de la crête explique respectivement 29,62 %, 26,38 % et 44,45 % de la variabilité du poids pour les espèces Adélie, Chinstrap et Gentoo.

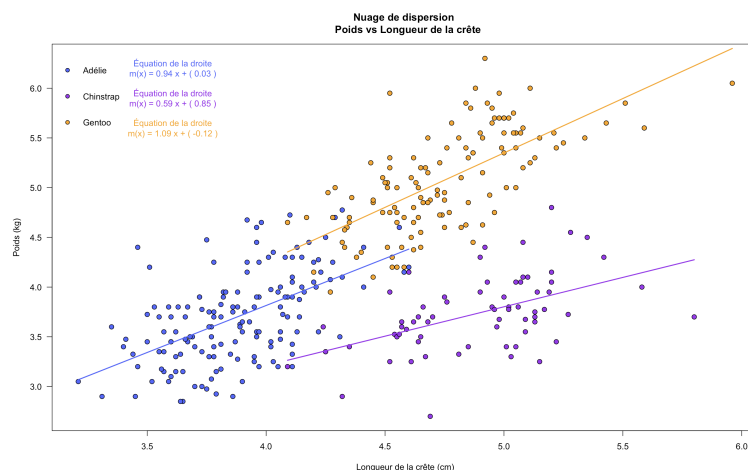


Fig. 9. Le poids des manchots en fonction de la longueur de la crête supérieur du bec, selon l'espèce

3.3 Le poids des manchots en fonction de la profondeur de la crête supérieur du bec, selon l'espèce

Ce dernier graphique (cf. Fig. 10) nous permet de mieux comprendre les études 1.3 et 2.3 (cf. sections 1.3 & 2.3). Il met en évidence les "deux nuages" qui sont en réalité trois. Ce n'est pas le sexe qui a influencé nos études, mais l'espèce. Plus précisément, nous étions confrontés à un paradoxe de Simpson, où un phénomène observé dans plusieurs groupes s'inverse lorsque les groupes sont combinés. Après la décomposition des sous-ensembles (espèces), nous distinguons correctement les associations linéaires. Nos précédentes études étaient donc mal interprétées. Cette fois-ci, nous pouvons observer une association linéaire croissante entre les variables poids et profondeurs de la crête. Pour les espèces Adélie, Chinstrap et Gentoo nous obtenons respectivement ces coefficients de corrélation linéaire de Pearson 0.58, 0.6 et 0.72. Nous en concluons qu'il y a une association linéaire modérée entre les variables pour l'espèce Adélie, et une forte association linéaire positive pour les espèces Chinstrap et Gentoo. Pour chaque augmentation de 0.2 cm de la profondeur du bec, nous observons en moyenne une augmentation de 0.27 kg chez l'espèce Adélie, de 0.26 kg chez l'espèce Chinstrap et de 0.46 kg chez l'espèce Gentoo. Chez l'espèce Adélie, la variable de la profondeur de la crête explique 33.66 % de la variabilité du poids, chez l'espèce Chinstrap 36.54 % et chez l'espèce Gentoo 52.27 %.

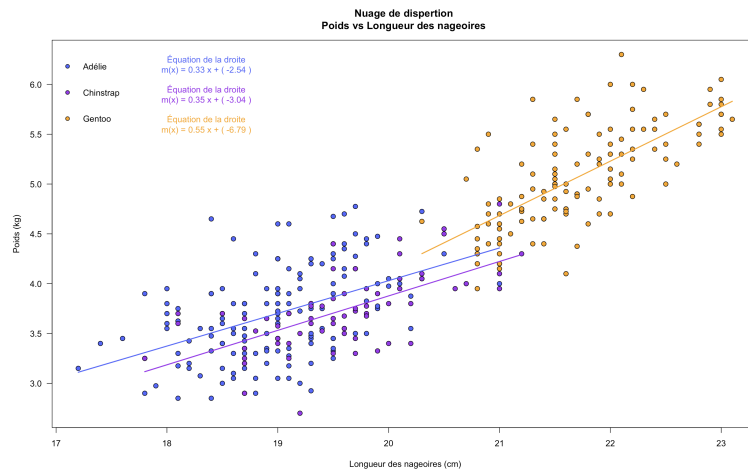


Fig. 10. Le poids des manchots en fonction de la profondeur de la crête supérieur du bec, selon l'espèce

Conclusion

À travers ce rapport, nous proposons une analyse du poids des manchots des archipels Palmer en Antarctique. Cette étude, divisée en trois parties, nous permet de mieux visualiser et comprendre l'évolution du poids des manchots en fonction de différents paramètres.

La première partie, met en avant que la variable de la longueur des nageoires est celle qui explique le mieux les variations du poids des manchots. De plus, nous avons observé que le modèle de régression proposé sur la figure 4 comparant l'évolution du poids en fonction de la profondeur du bec (cf. Fig. 4), n'était pas le plus approprié du fait de l'association linéaire décroissante.

La seconde partie nous permet d'étudier le poids des manchots selon leur sexe. Cette étude montre que le sexe ne fait pas vraiment varier le coefficient directeur de la droite des moindres carrées, ce qui signifie que l'évolution du poids est similaire quel que soit le sexe. En revanche, l'ordonnée à l'origine diffère, car les mâles sont plus lourds que les femelles.

La dernière partie portait sur l'étude du poids selon les espèces. Ici nous avons pu voir que l'espèce du manchot affecte fortement le poids des manchots. De plus, cette étude nous a permis de corriger les erreurs dues au paradoxe de Simpson présentes dans l'étude de la variable poids en fonction de la profondeur du bec.

4 Annexe

Ci-dessous deux image pour mieux visualiser les parties du corps étudiés des manchots (cf. Fig. 11 & 12) :

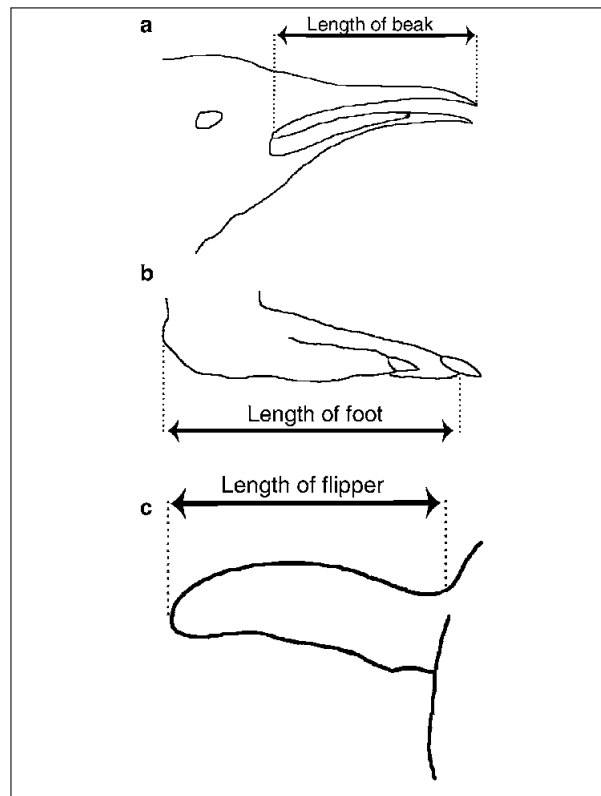


Fig. 11. Les mesures du manchots 1

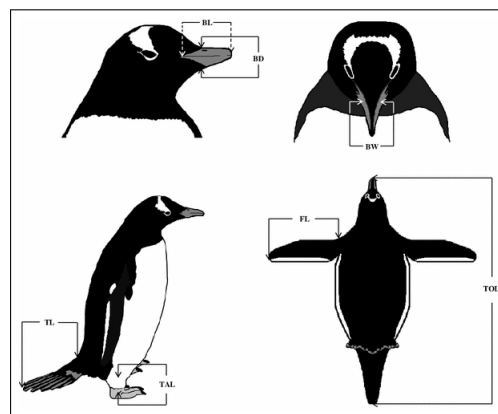


Fig. 12. Les mesures du manchots 2

Table des figures

1	Variables du jeu de données	3
2	Le poids des manchots en fonction de la longueur des nageoires	4
3	Le poids des manchots en fonction de la longueur de la crête supérieur du bec	4
4	Le poids des manchots en fonction de la profondeur de la crête supérieur du bec	5
5	Le poids des manchots en fonction de la longueur des nageoires, selon le sexe	6
6	Le poids des manchots en fonction de la longueur de la crête supérieur du bec, selon le sexe	6
7	Le poids des manchots en fonction de la profondeur de la crête, selon le sexe	7
8	Le poids des manchots en fonction de la longueur des nageoires, selon l'espèce	8
9	Le poids des manchots en fonction de la longueur de la crête supérieur du bec, selon l'espèce	8
10	Le poids des manchots en fonction de la profondeur de la crête supérieur du bec, selon l'espèce	9
11	Les mesures du manchots 1	11
12	Les mesures du manchots 2	11