

Stochastik Projekt

Table of Contents

Vorgehensweise.....	2
Grundfragen & Hypothesen.....	3
Wie sind Männchen und Weibchen verteilt	3
Auf welcher Insel leben die meisten Pinguine.....	5
Wie sieht die Aufteilung auf die Rassen aus	6
Wie sind die Rassen auf die Inseln verteilt	6
Wie sehr hängt das Geschlecht mit der Ausprägung der Merkmale zusammen	8
Männchen sind etwas größer als Weibchen.....	8
Die Merkmale bei Männchen sind etwas besser ausgeprägt als bei Weibchen.....	10
Größere Masse bedeutet stärker ausgeprägte Merkmale	15
Korrelieren Schnabeltiefe und Schnabellänge.....	17
Beeinflusst Rasse die Merkmale	18
Schnabellänge zu Rasse	18
Schnabeltiefe zu Rasse	20
Flügelänge zu Rasse	22

Vorgehensweise

Als erstes haben wir uns auf die Programmiersprache Python festgelegt, dies geschah schon vor dem Aussuchen des Datensatzes.

Danach haben wir uns Grundsätzliche Fragen gestellt, wie etwa, wie sind Männchen und Weibchen verteilt und wie sind die Rassen auf die verschiedenen Inseln verteilt.

Nachdem wir dieses analysiert und beschrieben hatten, gingen wir über zu neuen Hypothesen über, zuerst haben wir Korrelationen von Geschlecht zu Merkmalausprägung aufgestellt, allerdings ist uns da aufgefallen, dass Korrelationsdiagramme nicht dafür geeignet sind und sind daher auf Boxplots umgestiegen für diese Art von Analysen. Dann haben wir als Unterstützung hierfür noch extra Daten wie der genaue Mittelwert ausgerechnet.

Nach diesem haben wir dann dieselbe Vorgehensweise für die Vergleiche von Rassen und Merkmalen. Als letztes gingen wir dann über die Korrelationen von Schnabellänge zu Tiefe und von Gewicht zu Flügelänge, hier haben wir dann tatsächlich uns dafür entschieden Korrelationsdiagramme zu nehmen, da dies wiederum schwer ist mit Boxplots dar zu stellen.

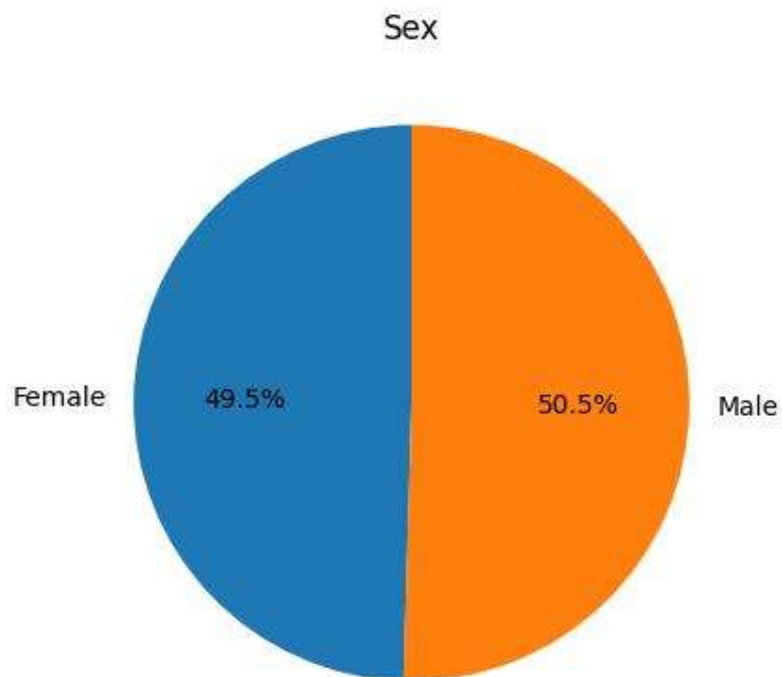
Schlussendlich wollten wir noch Gewicht zu Rassen Korrelation darstellen, allerdings, da Flügel und Gewicht sehr stark korrelieren und Rassen stark die Merkmale beeinflussen, konnten wir schon im Vorhinein feststellen, dass das Gewicht von Rasse zu Rasse variieren wird.

Grundfragen & Hypothesen

- Wie sind Männchen und Weibchen verteilt
 - Sie sind gleichverteilt
- Auf welcher Insel leben die meisten Pinguine
- Wie sieht die Aufteilung auf die Rassen aus
- Wie sind die Rassen auf die Inseln verteilt
- Wie sehr hängt das Geschlecht mit den Ausprägungen der Merkmale zusammen
 - Männchen sind etwas größer als Weibchen
 - Die Merkmale bei Männchen sind etwas besser ausgeprägt als bei Weibchen (Größe des Schnabels etc.)
- Größere Masse bedeutet stärker ausgeprägte Merkmale
- Korrelieren Schnabellänge und Schnabeltiefe
- Beeinflusst die Rasse die Merkmale

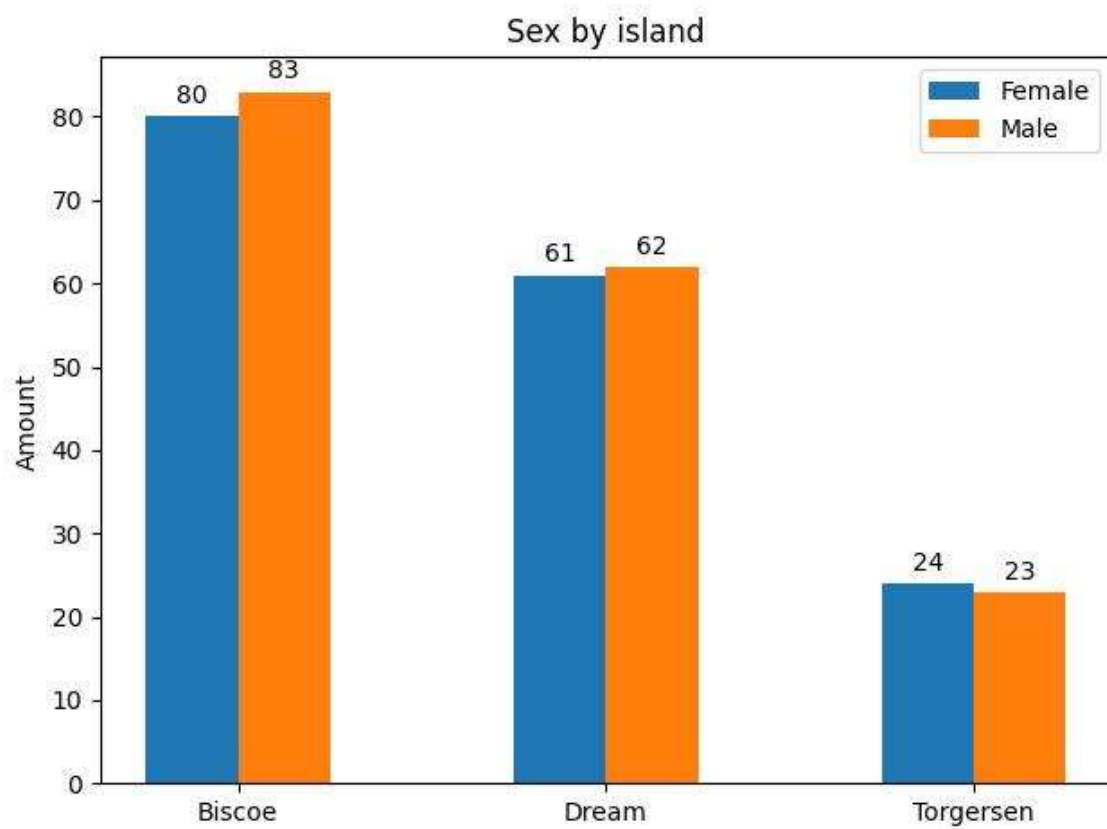
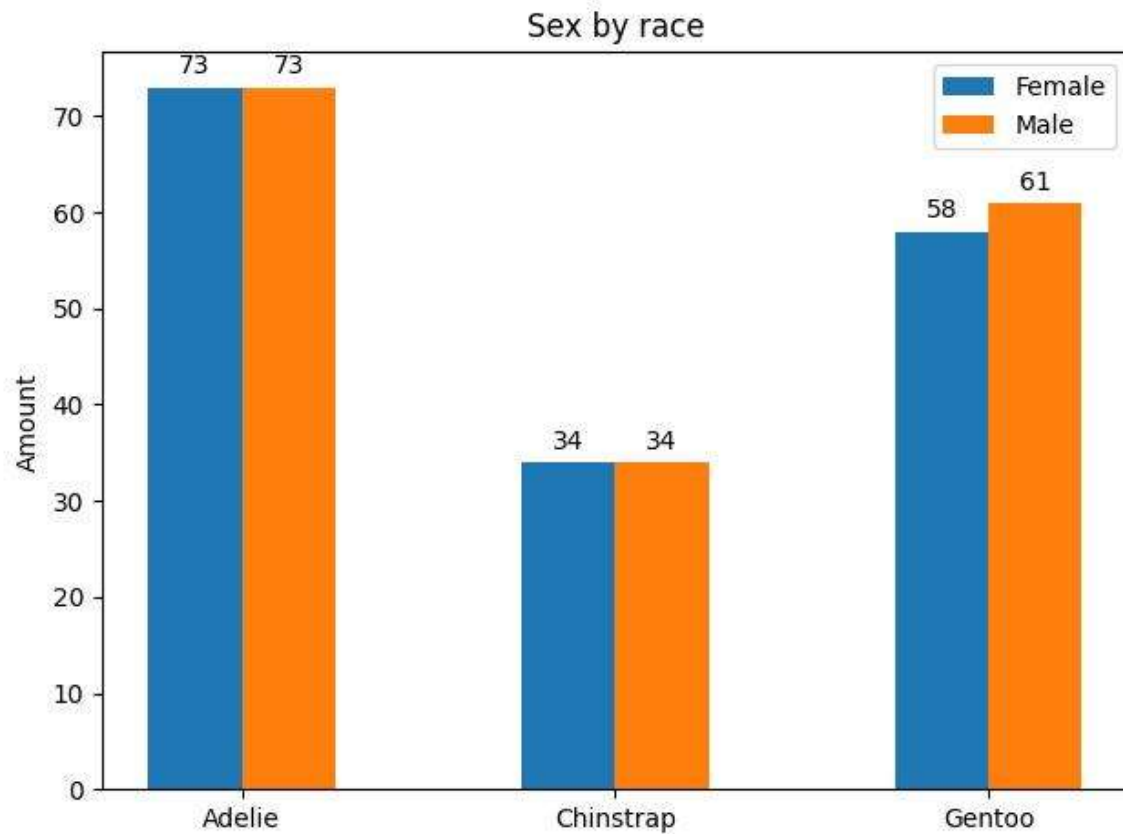
Wie sind Männchen und Weibchen verteilt

Die erste Leitfrage und meist auch die Grundlegendste ist wie die Verteilung von Männchen und Weibchen sich ausprägt.



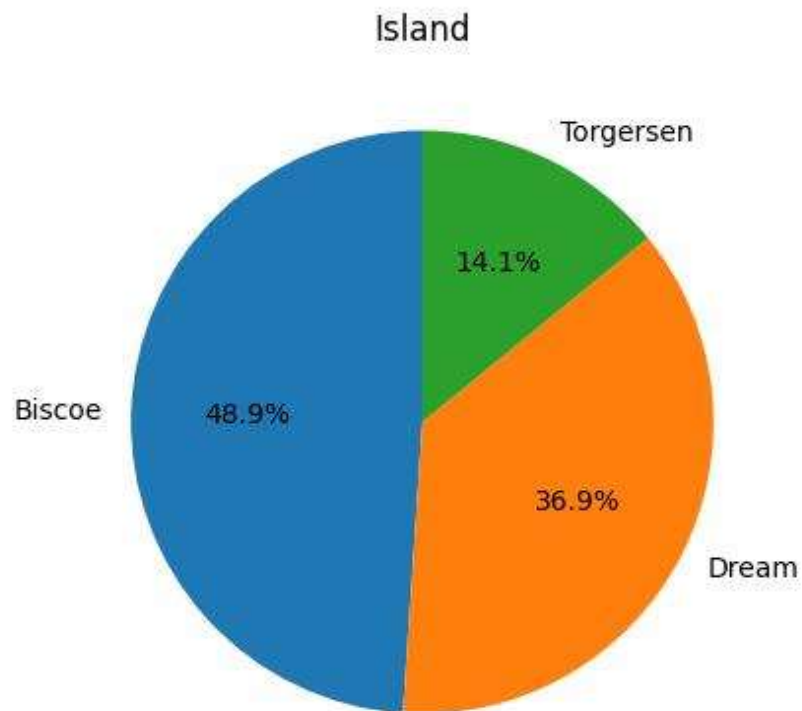
Und wie Anhand des Pie-Chart erkannt werden kann, sind die Geschlechter annähernd Gleichverteilt mit 49,5% Weibchen und 50,5% Männchen.

Als Nächstes liegt nahe, ob diese Verteilung weiterhin so gerecht bleibt über alle Rassen und Inseln hinweg.



Wie Anhand der 2 Charts gesehen werden kann, ist dass die Geschlechterverteilung auf den Inseln und in den Jeweiligen Rassen, immer annähernd den 50% entsprechen.

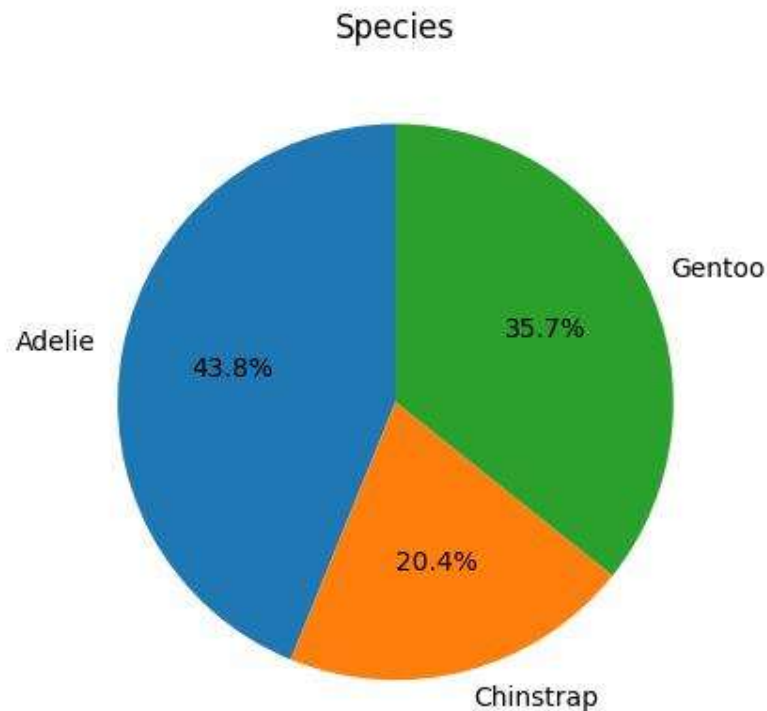
Auf welcher Insel leben die meisten Pinguine



Etwa 50% der Pinguine leben auf der Insel Biscoe, während etwas mehr als ein drittel auf Dream leben. Auf Torgersen leben hingegen dann nur noch 14,1% der Gesamtheit, daher sind die allgemein-Daten für Torgersen etwas weniger aussagend, da sie selbst nicht viel beitragen.

Wie sieht die Aufteilung auf die Rassen aus

Eine weitere Grundlegende Frage, welche sich stellt, ist wie die Verteilung auf die 3 Rassen aussehen könnte.

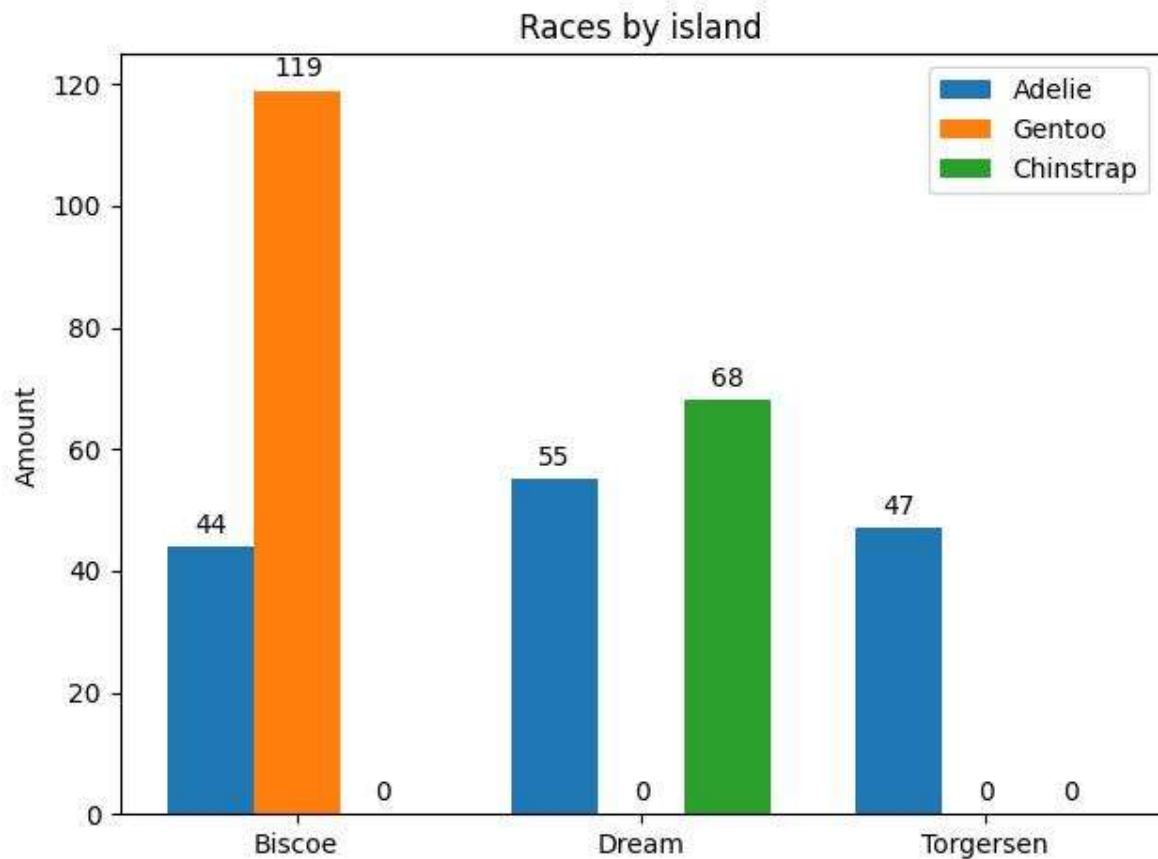


Diese Verteilung ähnelt sich relativ stark der Aufteilung auf den Inseln, allerdings mit leichten Abweichungen.

Die Rasse Adelie ist die vorherrschende mit 43.8%, knapp dahinter mit leicht mehr als ein Drittel befindet sich Gentoo. Mit ganzen 15% weniger Als Gentoo hat Chinstrap dann mit Abstand die wenigsten Individuen zu verzeichnen.

Wie sind die Rassen auf die Inseln verteilt

Als nächstes liegt dann natürlich die Frage nahe, wie die Rassen auf die verschiedenen Inseln verteilt sind, da ausgegangen werden kann, dass nicht jede Insel nur eine Rasse beinhalten kann, da die Zahlen der 2 vorhergehenden Auswertungen unterschiedlich ausfallen.



Wie erwartet sind die Rassen nicht 1 zu 1 auf die Inseln verteilt, spricht nicht 100% jeder Rasse sind auf der jeweiligen Insel. Was jedoch zu 100% nur auf einer Insel sind, sind die 119 Gentoo Pinguine, welche nur auf Biscoe sich befinden. Allerdings teilen die sich die Insel mit 44 weiteren Adelie Pinguine, ungefähr ein Viertel deren Gesamtbestandes. Auf der Dream Insel ist die Verteilung zwischen Adelie und Chinstrap annähernd 40 zu 60, mit 55 Adelie und 68 Chinstrap Pinguinen. Die letzte Insel Torgersen ist dann richtig Speziell, da sie ungefähr ein Drittel der Adelie beherbergt, allerdings machen diese auch 100% der Bevölkerung auf der Insel aus.

Durch diese Erkenntnis ist es relativ sinnbefreit weiters nach Inseln zu analysieren, da sich teilweise mehrere Rassen auf einer Insel befinden oder die Rasse ist auf mehrere Verteilt, daher sind diese Analysen nicht sehr aussagekräftig.

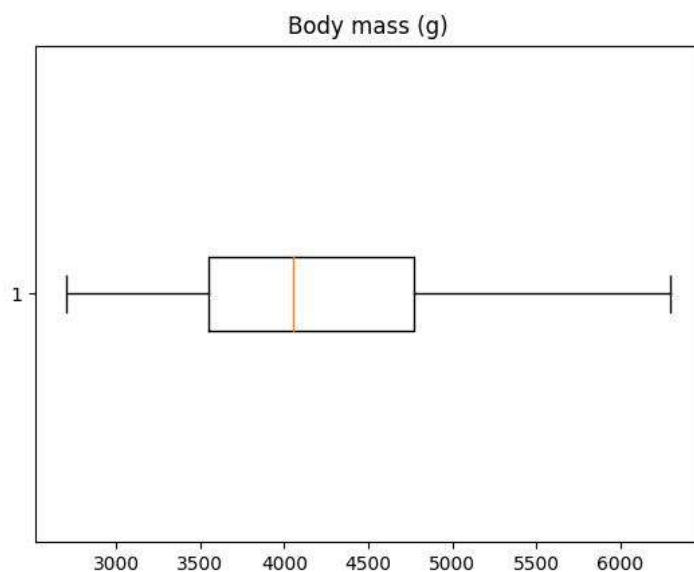
Wie sehr hängt das Geschlecht mit der Ausprägung der Merkmale zusammen

Nun gibt es eine interessante Frage, welche mehrere Auswertungen voraussetzt. Einige Hypothesen welche aufgrund von standardmäßigen Naturgesetzen gesetzt werden könnten wären:

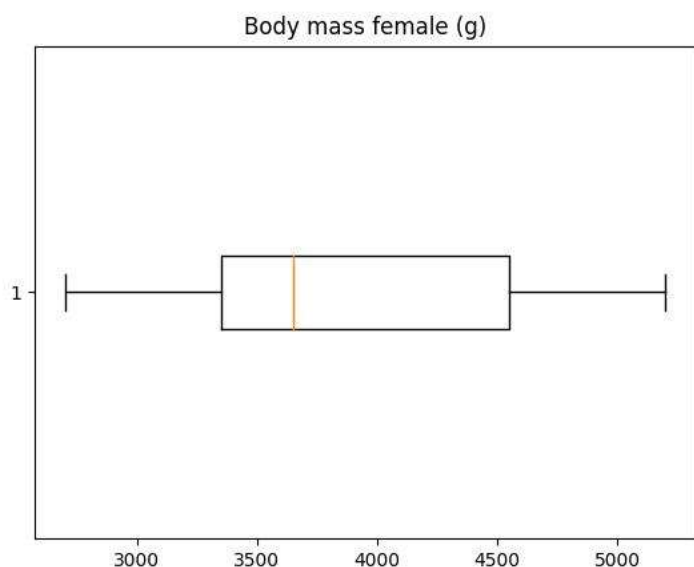
1. Männchen sind etwas größer als Weibchen
2. Die Merkmale bei Männchen sind etwas besser ausgeprägt als bei Weibchen (Größe des Schnabels etc.)

Männchen sind etwas größer als Weibchen

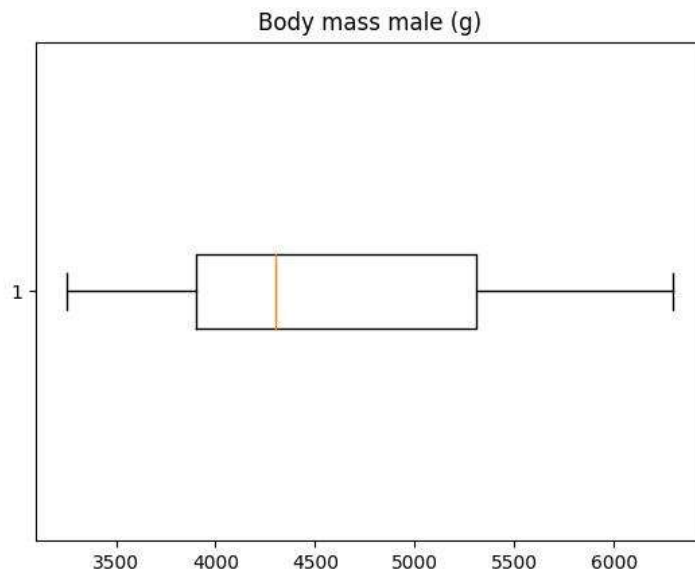
Hierfür wäre es möglich gewesen, Korrelationsdiagramme zu verwenden, allerdings sind diese relativ unattraktiv bei nur 2 Werten, daher werden hier Boxplots verwendet.



Der allgemeine Durchschnitt der Körper-masse liegt, unabhängig vom Geschlecht, bei 4207,05 Gramm. Wobei mehr als die Hälfte zwischen 3500 und 5000 Gramm liegen.



Bei Weibchen zieht sich das dann schon etwas mehr nach Unten, der Durchschnitt bei denen liegt bei 3862,27 Gramm, 320 Gramm oder 8% unter dem Allgemein-Schnitt. Zudem zeigt sich anhand des Boxplots, dass kein Weibchen über 5500 Gramm kommt, im Gegensatz zur Allgemeinheit. Somit steht schon fest, dass Männchen dies irgendwie ausgleichen müssen, um auf den Allgemein-Schnitt kommen zu können.

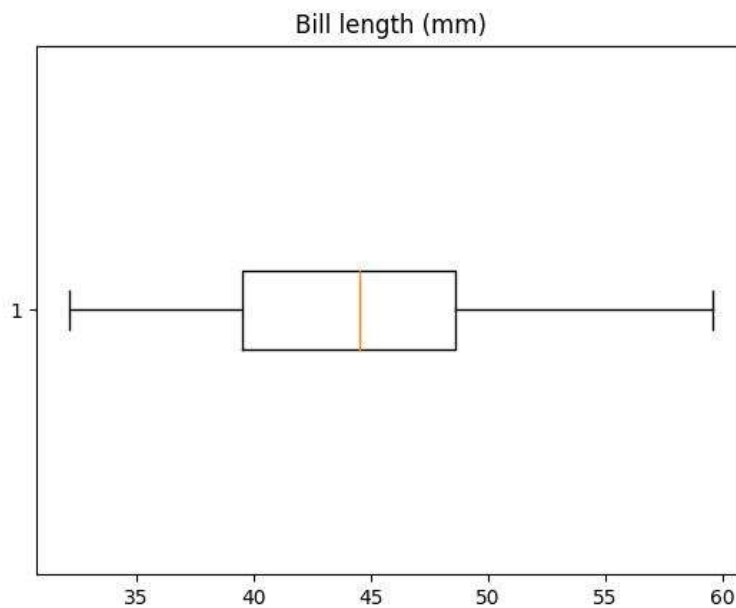


Die Männchen sind dann, wie erwartet im Durchschnitt schwerer mit 4545,68 Gramm. Das sind 18,8% mehr als die Weibchen. Zudem haben die Männchen noch ein höheres Minimum und Maximum beim Gewicht

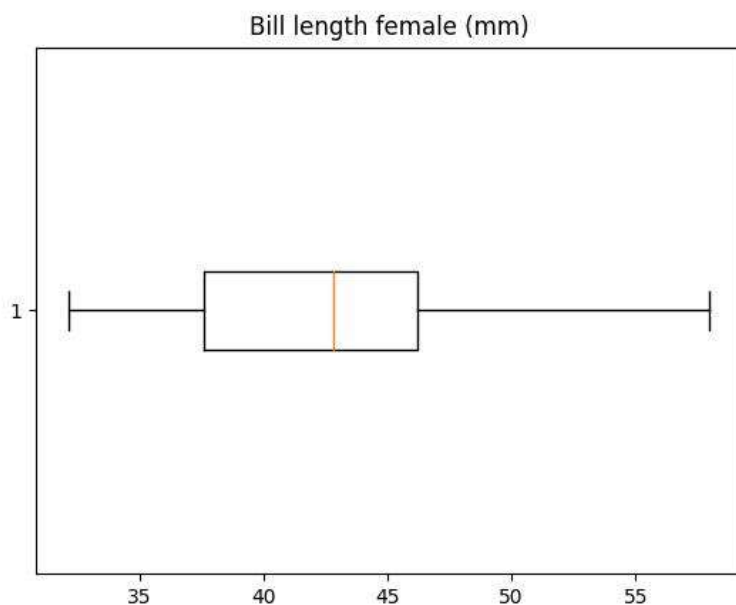
Somit ist die Hypothese „Männchen sind etwas größer als Weibchen“ hier bestätigt

Die Merkmale bei Männchen sind etwas besser ausgeprägt als bei Weibchen

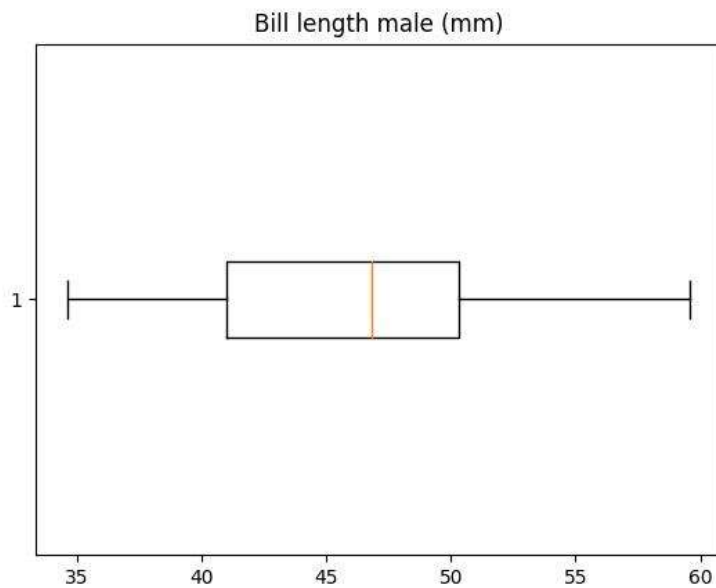
Hier gilt dasselbe wie bei der oberen Hypothese, Boxplots sind angenehmer bei einzelnen Werten.



Angefangen mit der Schnabellänge, diese beträgt im Durchschnitt 43,99mm, wobei mehr als 50% zwischen 40 und 50mm liegen.

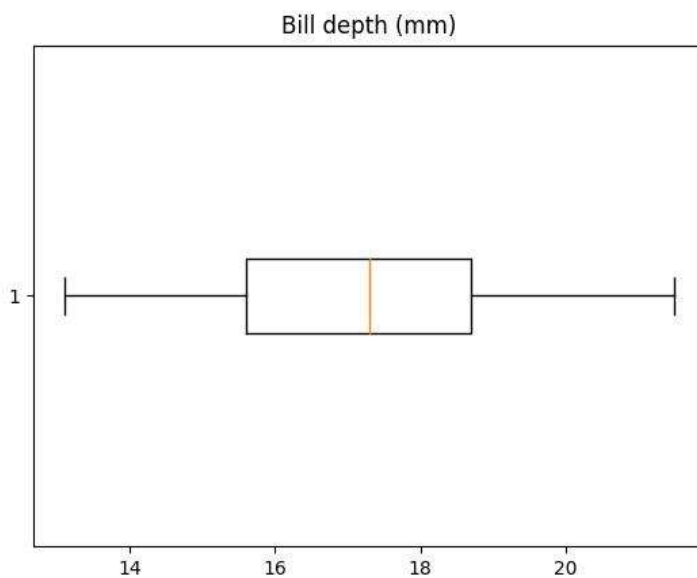


Bei Weibchen ist diese im Durchschnitt bei 42,09mm, das sind 1,9mm weniger als der Allgemeinschnitt. Zudem befinden sich die 50% etwas weiter unten, zwischen ungefähr 37 und 46mm. Und auch das Maximum ist niedriger mit etwa 58.

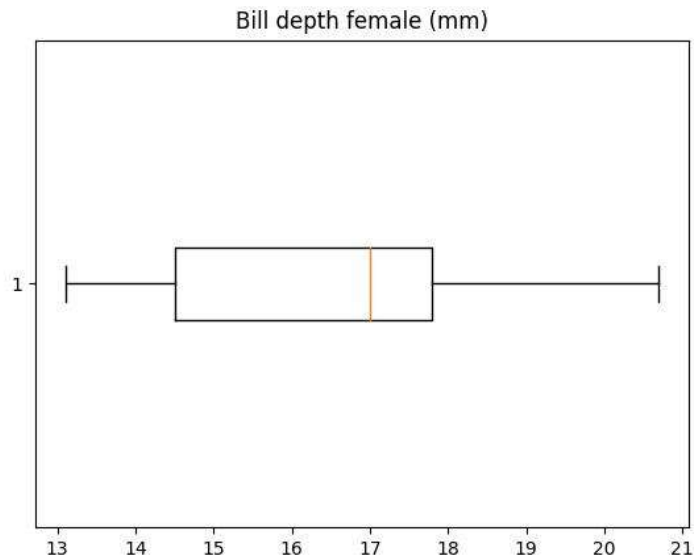


Bei Männchen liegt der Durchschnitt der Schnabellänge bei 45.85mm, was 1,86mm mehr sind als beim Allgemein-Schnitt und 8% mehr als bei Weibchen. Zudem ist der Minimal- und Maximalwert etwas angehoben.

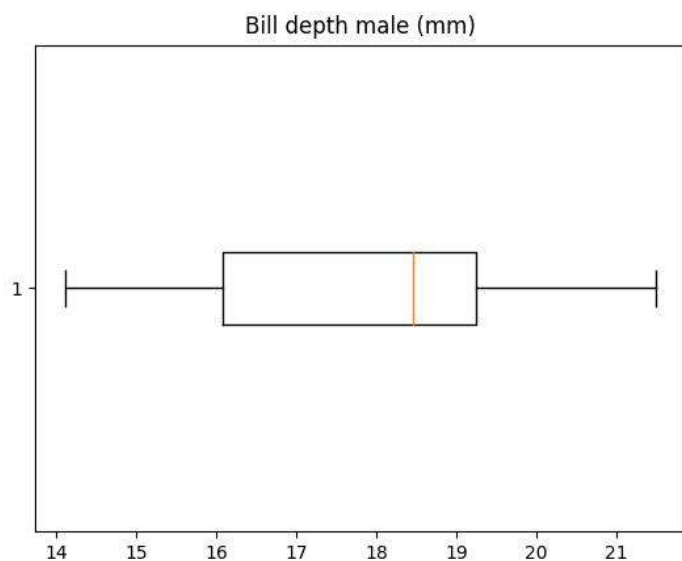
Somit steht fest, Schnäbel sind bei Männchen etwas länger als bei Weibchen. Jetzt stellt sich die Frage, ob dies auch bei der Tiefe der Fall ist.



Die Tiefe beträgt im Durchschnitt 17,16mm, wobei sich mehr als 50% zwischen 15,5 und 19mm befinden.



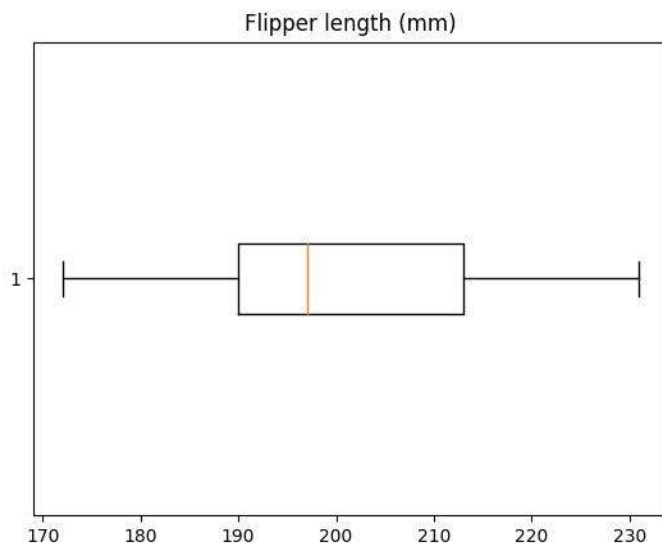
Bei Weibchen befindet der Durchschnitt bei 16,42mm Tiefe, was 0,74mm unter dem Allgemeinschnitt liegt. Auch hier liegen wieder Min und Max werte leicht unter dem Allgemeinschnitt, allerdings ist dies hier etwas weniger extrem wie bei vorherigen Beispielen.



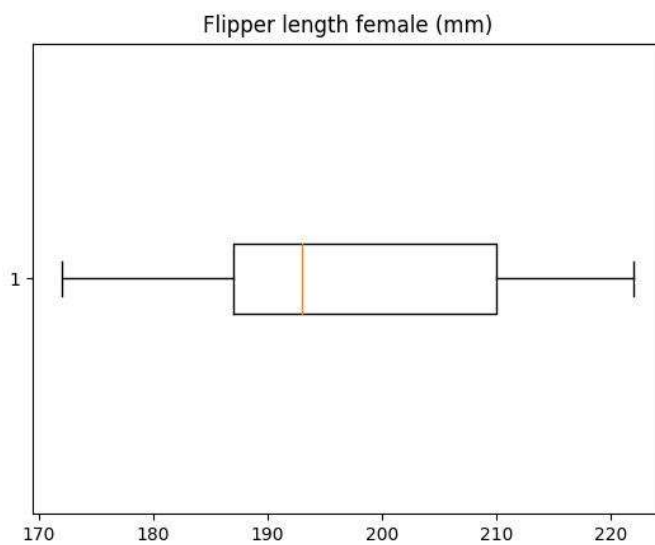
Männchen haben im Durchschnitt eine Tiefe des Schnabels von 17,89mm. Dies sind wieder 9% mehr als bei Weibchen. Auch ist das Minimum bei 14 anstatt bei 13 und das Maximum befindet sich über 21,5 im Gegensatz zu unter 21mm bei Weibchen.

Somit steht auch fest, dass Männchen einen tieferen Schnabel besitzen.

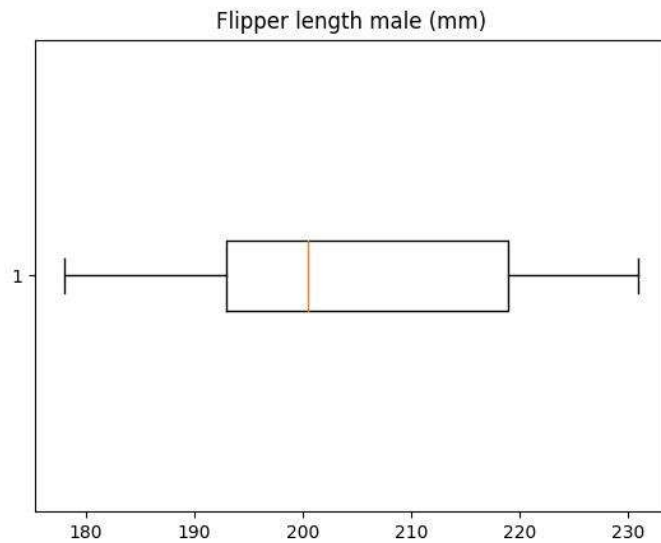
Jetzt steht die Frage ob dies auch bei den Flügeln ähnlich aussieht.



Allgemein liegt der Durchschnitt bei 200,97mm Länge, mit einem Minimum von ungefähr 170 und Maximum bei 231mm. Ca. 50% der Pinguine haben hierbei auch eine Flügellänge zwischen 190 und 213mm.



Weibchen haben eine Durchschnittliche Flügellänge von 197,36mm was wiederum 3,61mm unter dem Allgemein-Schnitt liegt. Auch der Maximalwert ist geringer.

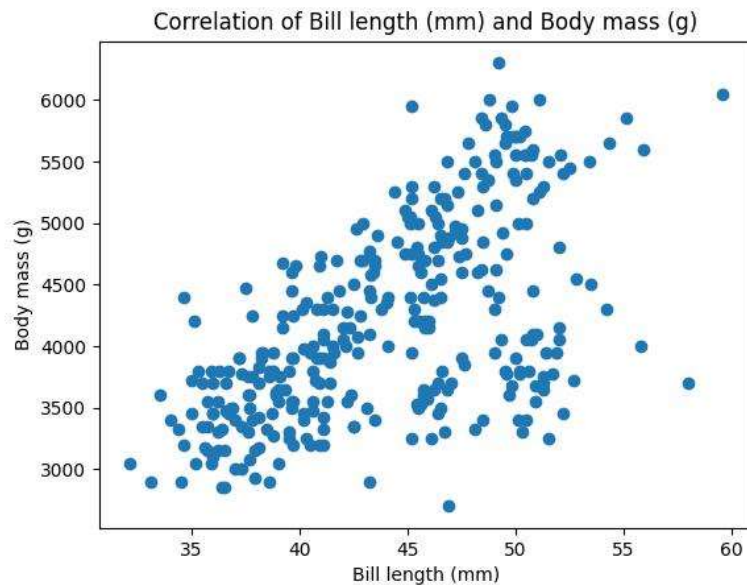


Männchen haben dann wie erwartet eine höhere durchschnittliche Flügellänge als Weibchen mit 204,51mm. Dies sind 3,6% mehr als bei Weibchen, dieses Merkmal hat somit die kleinste Abweichung von Männchen zu Weibchen von den 3 vorhandenen.

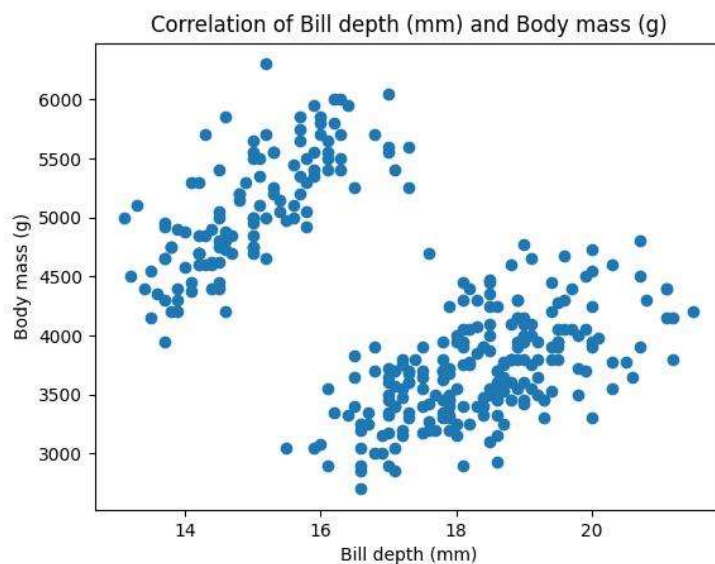
Hiermit kann mit Sicherheit gesagt werden, dass das Geschlecht einen Einfluss hat auf die Ausprägung auf Schnabel und Flügel. Der Schnabel ist bei Männchen im Durchschnitt 8% Länger und Tiefer als bei Weibchen und auch die Flügel sind 3,6% länger.

Größere Masse bedeutet stärker ausgeprägte Merkmale

Nachdem der Vergleich zwischen Männchen und Weibchen abgeschlossen ist, ist bekannt, dass Männchen stärkere Ausprägungen der Merkmale besitzen und auch schwerer sind. Allerdings stellt sich jetzt eine weitere Frage: Hat das Gewicht einen Einfluss auf die Ausprägungen?

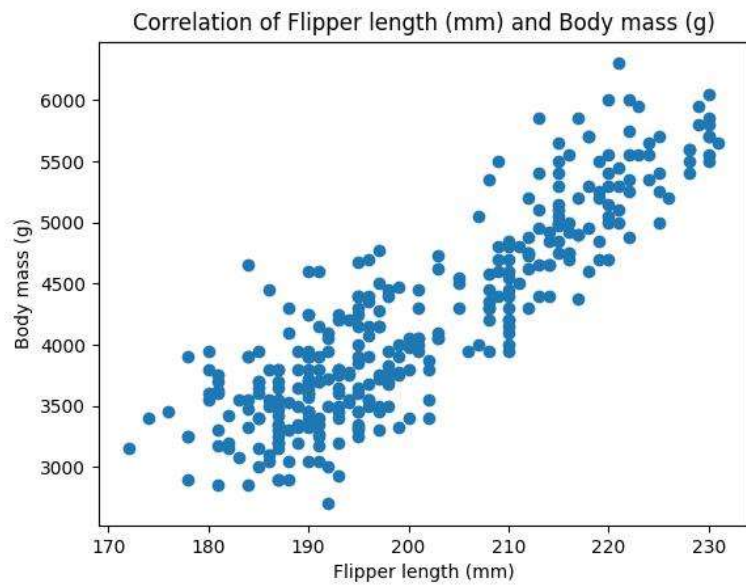


Die Körper-Masse und Schnabellänge ist die erste Korrelation die untersucht wird und wie im Diagramm gesehen werden kann, besteht eine mittel starke Korrelation zwischen den beiden Werten. Der Koeffizient liegt hierbei 0,59.



Bei der Korrelation von Schnabeltiefe sieht das Diagramm zwar drastisch anders aus, allerdings besteht hier wieder eine mittlere Korrelation mit einem Koeffizienten von 0,47.

Somit steht fest, dass die Ausprägungen des Schnabels zwar mit dem Körpergewicht zusammenhängen, aber auch stark abweichen können und nicht daran 100% gebunden sind.

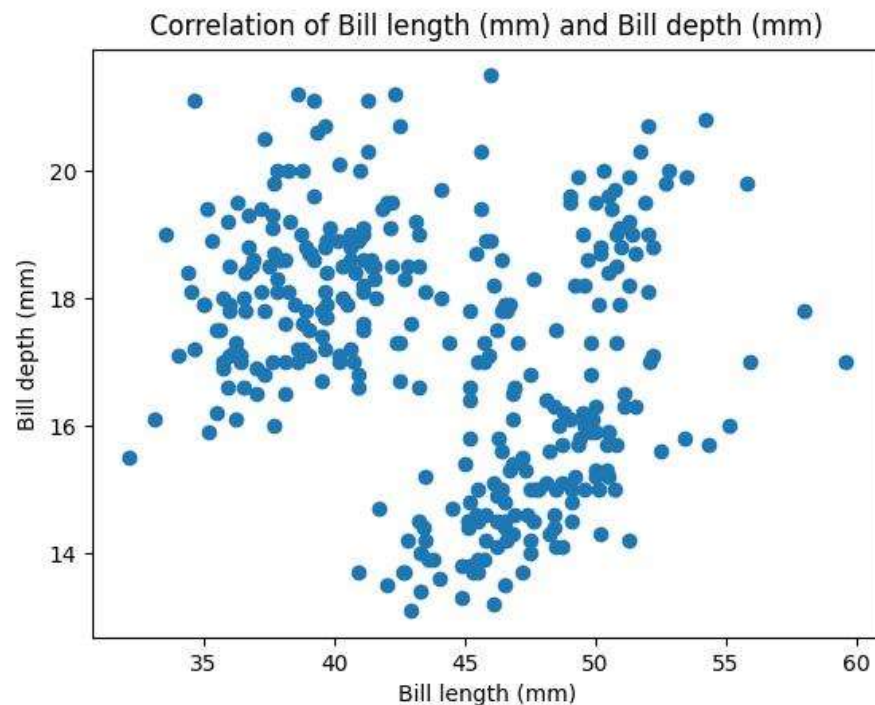


Bei der Flügelänge sieht dies jedoch schon ganz anders aus, hier kann schon anhand des Diagrammes erkannt werden, dass eine starke Verbindung zwischen Gewicht und Flügelänge vorherrscht. Auch der Koeffizient von 0,87 belegt diese Hypothese.

Somit steht fest, dass Gewicht einen Einfluss hat auf Ausprägungen. Auf den Schnabel etwas weniger als auf die Flügel, aber die Verbindung steht bei beiden Merkmalen.

Korrelieren Schnabeltiefe und Schnabellänge

Eine weitere Hypothese, die gestellt werden könnte, ist: Ein längerer Schnabel bedeutet auch ein tieferer Schnabel.



So logisch diese Aussage auch zu scheinen mag, ist sie nichtzutreffend. Schon der Graph zeigt eine Diskrepanz der beiden Werte auf, viele lange Schnäbel sind nicht sonderlich tief und auch anders herum. Der Korrelationskoeffizient befindet sich auch nur bei -0,23, was bedeutet dass so gut wie keine Verbindung zwischen Länge und Tiefe vorherrscht.

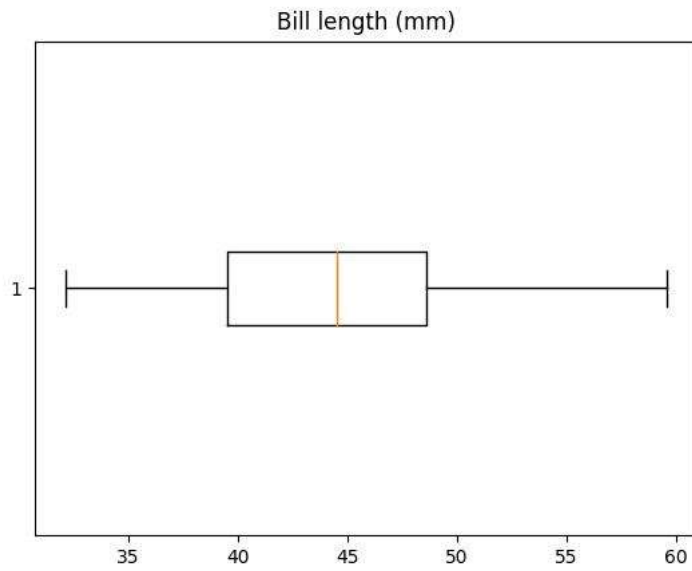
Somit ist die Hypothese „Längerer Schnabel bedeutet tieferer Schnabel“ eindeutig nicht zutreffend.

Beeinflusst Rasse die Merkmale

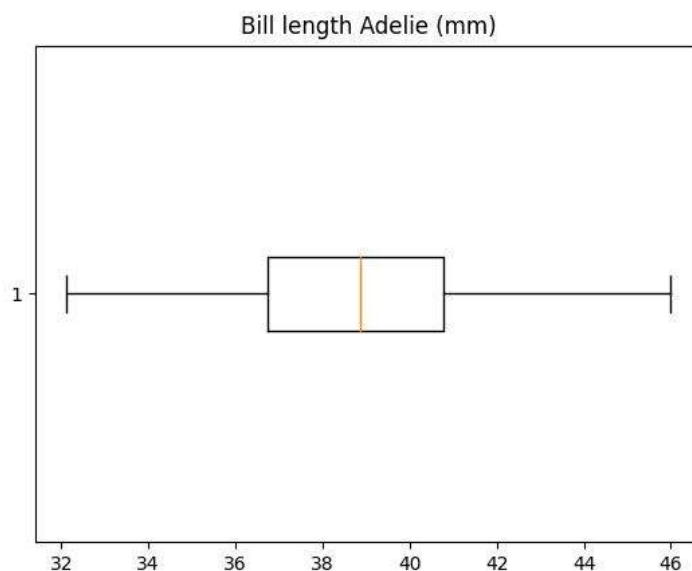
Da relativ früh schon belegt wurde, dass nach Inseln relativ Sinnbefreit ist, ist nur noch eine Hypothese übrig, welche untersucht werden sollte: Die Rasse beeinflusst die Ausprägungen der Merkmale.

Auch hier gilt wieder, dass Boxplots sinnvoller sind als Korrelationsdiagramme, da auch mit 3 werten die Streuung nicht schön wird.

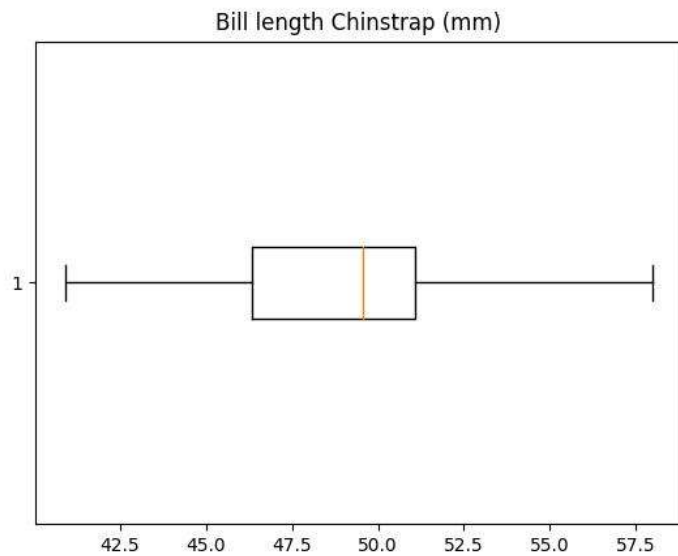
Schnabellänge zu Rasse



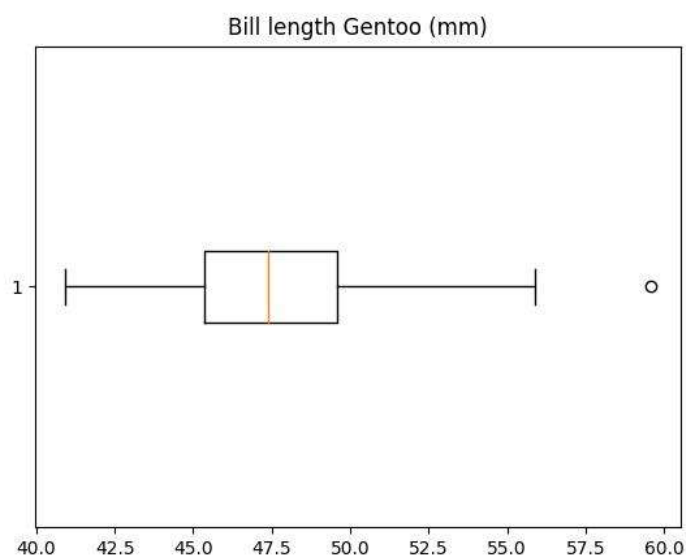
Die Durchschnittliche Länge beträgt hier wieder 44mm.



Die Adelie haben im Gegenzug einen weitaus kürzeren Schnabel mit nur 38,82mm, was nur 11,78% unter dem Durchschnitt liegt.



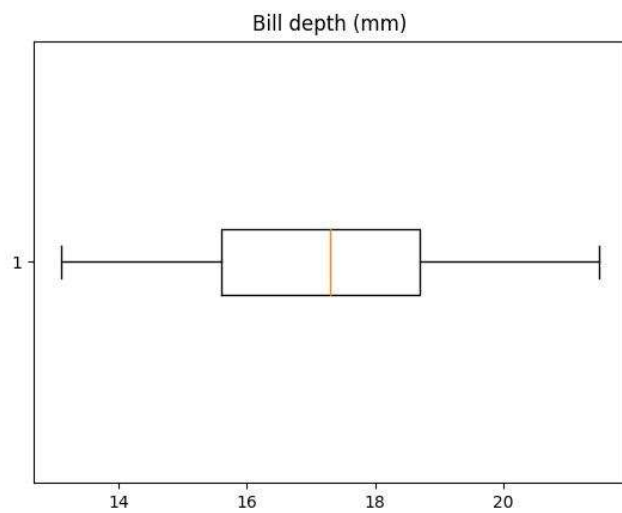
Bei den Chinstrap sieht dies wieder drastisch anders aus, es hebt sich hier so stark von den Adelie werten ab, dass der Minimalwert der Chinstrap über dem Mittelwert der Adelie liegt. Der durchschnittliche Chinstrap Pinguin hat hierbei eine Schnabellänge von 48,83mm, dies sind ganze 10mm mehr als bei den Adelie. Allerdings liegt das Maximum von 57,7mm unter dem Allgemein-Maximum von 60mm.



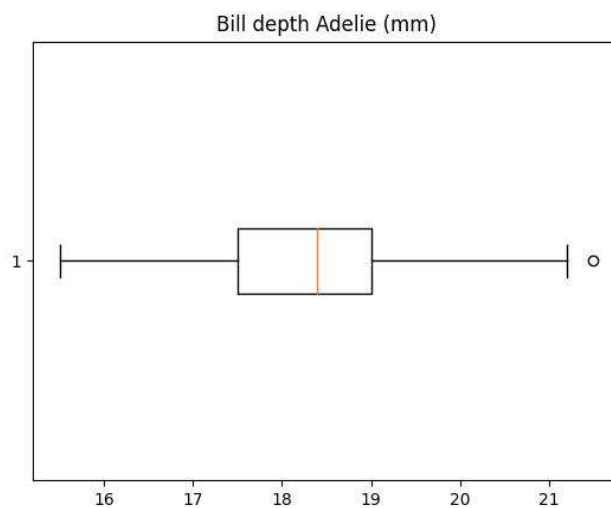
Die Gentoo Rasse hat eine durchschnittliche Schnabellänge von 47,57mm, welches 1,46mm unter der Durchschnittslänge der Chinstrap-Schnäbel liegt. Auch hier ist das Minimum höher als der Durchschnitt der Adelie, somit kann mit hoher Sicherheit gesagt werden, dass die Adelie mit Abstand die kleinsten Schnäbel haben. Allerdings sind Chinstrap und Gentoo sich relativ ähnlich von den Größen her, abgesehen von dem Ausreißer mit 60mm Schnabellänge bei den Gentoo.

Es kann also gesagt werden, dass Schnabellänge von den Rassen beeinflusst wird.

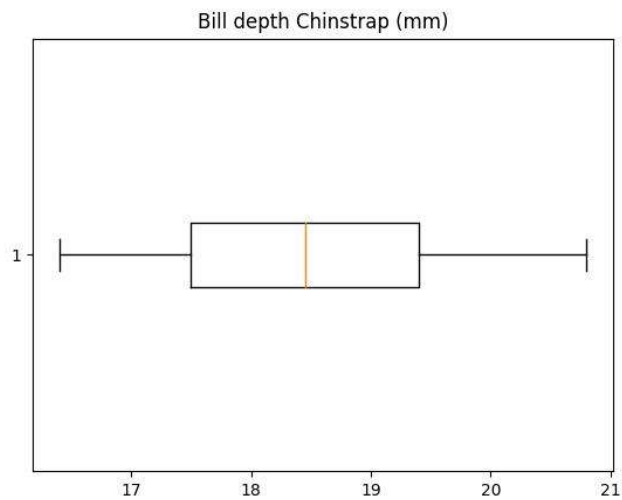
Schnabeltiefe zu Rasse



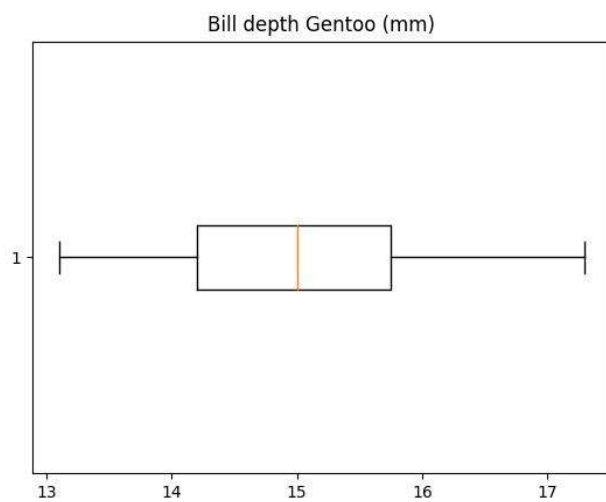
Der allgemeine Durchschnitt bei der Schnabeltiefe beträgt 17,16mm.



Die Adelie haben zwar kürzere Schnäbel als die Allgemeinheit, allerdings ist Ihre durchschnittliche Schnabeltiefe höher mit 18,35mm. Dies sind 7% mehr als der Allgemein-Schnitt. Die Adelie haben zusätzlich noch einen Ausreißer zu verzeichnen mit über 21,5mm.



Die Chinstrap haben eine ähnliche Durchschnittliche Tiefe wie die Adelle mit 18,83mm. Mit nur einem Unterschied von 0,48mm.

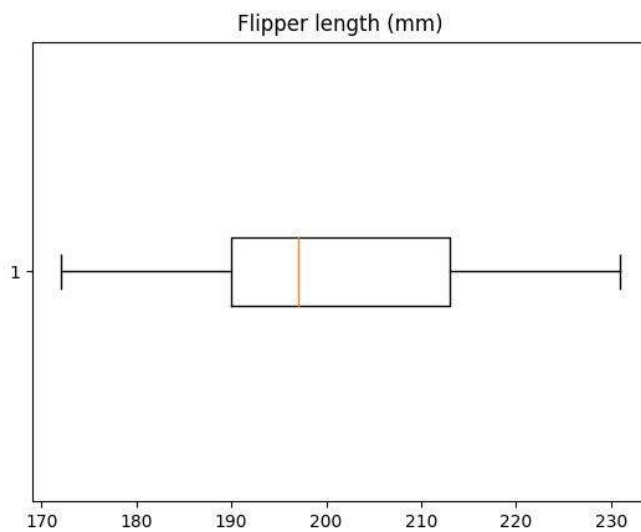


Die Gentoo haben in dem Vergleich mit Abstand den flachsten Schnabel mit einem Mittelwert von 14,99mm. Dies sind ganze 2,17mm weniger als der Allgemein-Schnitt.

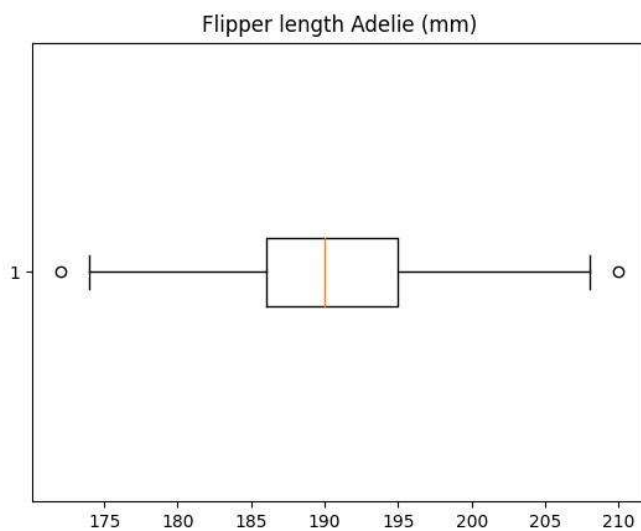
Somit kann gesagt werden, schon allein, weil die Gentoo weitaus flachere Schnäbel besitzen, dass Rassen einen Einfluss haben auf die Schnabeltiefe.

Flügelänge zu Rasse

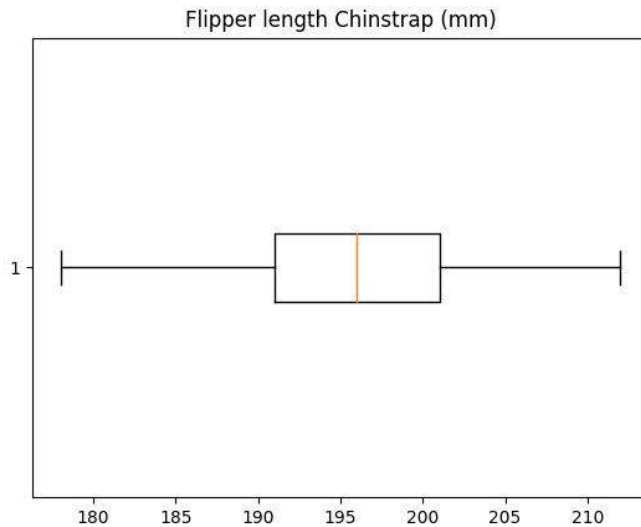
Als dritter Vergleich steht noch aus, ob die Rassen einen Unterschied in der Flügelänge machen.



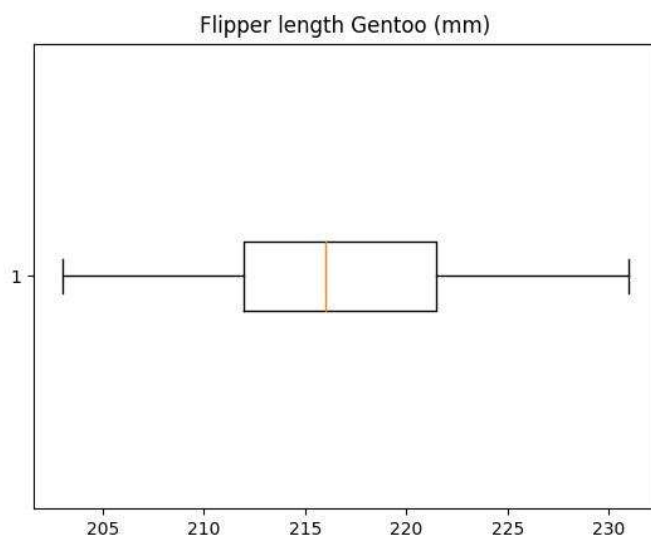
Die durchschnittliche Flügelänge ist hier wieder 200,97mm, wie bei vorherigen Beispielen.



Adelie Pinguine haben auch hier eine Ähnliche Tendenz wie bei dem Schnabellängenvergleich. Sie haben allgemein vergleichsweise kurze Flügel mit einem Durchschnitt von 190,1mm, dies sind 10,87mm weniger als beim Allgemein-Schnitt.



Auch die Chinstrap haben vergleichsweise kleine Flügelängen zu verzeichnen, mit nur 195,82mm sind sie nur 5,72mm länger als die kleinste Rasse Adelie.



Die Gentoo haben hier im Vergleich sehr lange Flügel. Sie besitzen eine durchschnittliche Flügelänge von 217,23mm. Mit diesem Wert sind sie ganze 16,26 über dem Schnitt.

Auch die Flügelänge ist somit stark mit den Rassen verbunden.

Fazit zu den Merkmalen ist somit, dass alle Merkmalsausprägungen stark damit zusammenhängen, von welcher Rasse die Pinguine sind.

Was Gewicht zu Rasse Korrelation angeht, kann dies schon von der Flügelänge abgeleitet werden, da Flügelänge und Gewicht zusammen eine sehr hohe Korrelation aufweisen, kann ziemlich sicher gesagt werden, dass Rassen auch mit dem Gewicht zusammenhängen.