PENGUINS : SEX AND SPECIES COMPARISON

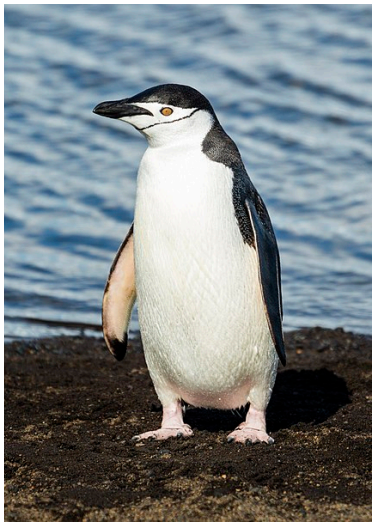
RELEVANT JMP PLATFORMS AND STATISTICAL TECHNIQUES

Graph Builder & Graphs :	Comparative Dotplots ; 3D Scatterplot
Multivariate :	Scatterplot Matrix
Fit Y by X :	One-Factor ANOVA
Distribution :	Univariate analyses, Confidence Intervals for the Population Mean, Tolerance Intervals for proportions of a Population.

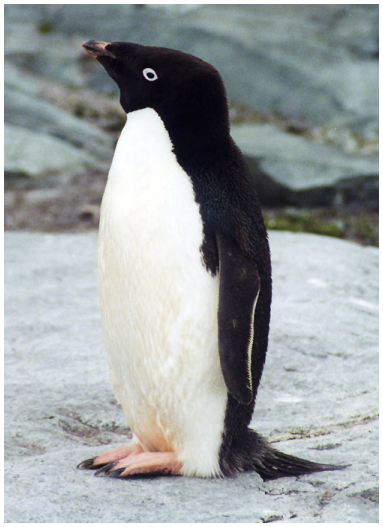
PROBLEM STATEMENT

To better understand similarities and differences in morphological features between males and females of different penguin species, a researcher from Simon Fraser University in collaboration with The Palmer Station Antarctica Long Term Ecological Research Network obtained measurements on a sample of penguins from three different species [1].

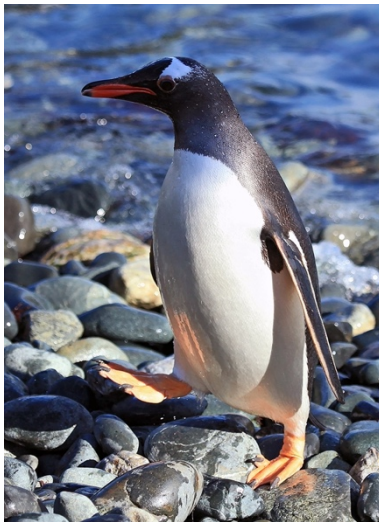
The dataset consists of measurements across 4 physcial features of 34 male and 34 female Chinstrap penguins, 71 male and 72 female Adélie penguins, and 61 male and 58 female Gentoo penguins.



Chinstrap



Adélie



Gentoo

DATA Penguins_Sex_and_Species_Comparison.jmp

Species	Species of Penguin (Chinstrap, Adélie,Gentoo)
Sex	Sex (Male or Female)
Culmen Length	Length (mm) of culmen*
Culmen Depth	Depth (mm) of culmen*
Flipper Length	Length (mm) of flipper
Body Mass	Body mass (grams)

** Culmen is the upper ridge of the beak.*

EXCERCISES

Tip: Before you begin the analyses, it would be best to create a row state using a unique color for each species and symbol for each sex. This can be done by selecting the two cells for a unique species/sex combination. Then choose "Rows > Row Selection > Select Matching Cells". Then with all the rows of a particular sex/species, Choose Rows > Colors" and "Rows > Markers" to create a unique row/symbol. Then repeat for the remaining species/sex combinations. Now all graphs created in any future analysis will utilize these row states.

1. Create comparative dotplots across species and sex for each of the 4 physical features. Describe the similarities and differences observed in these features.

2. Create a scatterplot matrix and parallel coordinate plot of the 4 features. Update any of your descriptions from Excerise 1 with any new insights observed in these visualizations.
   Tip: Create scatterplot matrix and parallel coordinate plot from the "Analyze > Multivariate Methods > Multivariate" platform. From within the report, select the red triangle to launch the Local Data Filter. Choose 'Species' and 'Sex' as the variables to use. Now specific Species and Sex combinations can be selected for comparison purposes.

3. Create a 3D Scatterplot matrix using the 4 features. Interactively explore these graphs by rotataing the axes. Update any of your descriptions created from Excerise 1 and 2 with any new insights observed in these visualizations.
   Tip: Create scatterplot matrix and

4. Perform a One Factor ANOVA for each feature using Sex as the X variable. Peform a second One Factor ANOVA for each feature using Species as the X variable. Identify any statistically significant differences observed in the means of each feature between the species and the sexes. Why are these analyses not the best ways to test/quantify statistical differences between the 2 sexes and between the 3 features?

5. For each Species, perform a separate One Factor ANOVA for each feature using Sex as the X variable. Describe any statistically significant differences in the features between the sexes within each species.
   Tip: Use Fit Y by X platform and choose Species as the 'By Variable'.

6. For each Sex, perform a separate One Factor ANOVA for each feature using Species as the X variable. Describe any statistically significant differences in the features between the species within each Sex.
   Tip: Use Fit Y by X platform and choose Sex as the 'By Variable'.
   Hint: Augment the analyses with a multiple comparisons to test/quantify all pairwise differences between the 3 species.

7. For each Species/Sex combination, produce a 95% Confidence Intervals for the population mean for each feature. Describe how one would interpret these intervals by demonstrating it for 2-3.
   Tip: Do so in the Analyze > Distribution platform, choosing both Species and Sex as the 'By Variable'. They'll be a total of 4 x 3 x 2 = 24 univariate analyses performed.

8. For each Species/Sex combination, produce a 99% Tolerance Interval (at 95% Confidence) for each feature. Describe how one would interpret these intervals by demonstrating it for 2-3.
   Tip: Tolerance Interval is found under the red triangle next the histograms.

SUPPLEMENTAL MATERIALS

1. Gorman KB, Williams TD, Fraser WR (2014) "Ecological Sexual Dimorphism and Environmental Variability within a Community of Antarctic Penguins (Genus *Pygoscelis*)". *PLoS ONE*, 9(3): e90081, 2014.
   https://doi.org/10.1371/journal.pone.0090081
   https://pallter.marine.rutgers.edu

Resources to learn more about these penguin species.

https://en.wikipedia.org/wiki/Pygoscelis

https://www.penguinworld.com/types/pygoscelis.html