

How many penguins?

Activity Summary This activity introduces children to the concepts of populations and samples. Using information obtained from a sample of “local” penguins, we infer how many penguins live in a colony in Antarctica.	
Activity Learning Outcomes: • Understand what the words “population” and “sample” mean. • Be able to use an estimate obtained from a sample to infer back to the population. • Appreciate the usefulness of this inference in real life applications.	Resources: • Aerial photograph of penguin guano on snow with appropriate scale (provided with this pack and online here) • “Choose Your Penguins!” handout (provided with this pack)
Video demonstration: • A video demonstrating a modified version of this activity is available on the RSS website at www.rss.org.uk/membership/faq/hands-on-statistics	

How to run the activity

- Begin the activity by asking the students if they can count how many penguins can be seen in the aerial photograph (they can’t)
- Once it has been established that it is impossible to count individual penguins in the photograph get the students to estimate the surface area of the guano using the provided scale (note that it may help to split the grid squares into smaller squares and count the number containing poo). The photograph shows approximately 150,000m² of guano – if necessary use the wording “Check you agree that the area of guano in the photograph is 150,000m²”.
- Encourage students to think about where they have seen penguins more conveniently than in Antarctica (zoo, aquarium etc.) and discuss how it may be easier to study these penguins than those in the wild.
- Introduce the students to the “local” penguins (on the choose your penguin handout) and explain that an estimate of the amount of guano produced per year per penguin can be found next to each one.
- Ask them each to select (sample) 3 penguins.
- Have the students calculate the mean amount of guano produced by their sampled “local” penguins using the formula:

$$\text{Mean amount of guano} = \frac{\text{Penguin 1 guano} + \text{Penguin 2 guano} + \text{Penguin 3 guano}}{3}$$

- Now have the students infer how many penguins are in the aerial photo, and hence in that colony within Antarctica using the formula:

$$\text{Number of penguins} = \frac{\text{Surface area of guano in the photo}}{\text{Mean amount of guano per penguin per year}}$$

Exploring the activity

- Why do all the students not have the same answer? Discuss the impacts of the variability in students' choices of penguins.
- What happens if the sample of penguins is not random? For example, what if the male penguins tend to do more poo than the females – did the students pick at random or were they biased towards one or the other?
- What happens if a larger sample of “local” penguins is used?
- How certain are we in our estimate of penguin poo area in the photo?
- What assumptions might we be making here? (e.g. this method assumes that no two penguins poo in the same place – will that bias our estimate?)
- Cross-curricular opportunities to explore: discuss estimation of animal populations, relating to keeping track of endangered species. Students might have come across other methods e.g. the RSPB big garden bird watch. The aerial photograph method couldn't be used with animals in jungles – what other methods might be used?

What's going on?

- A **population** is a theoretical concept used to describe the entire group. The population of interest in this activity is all the penguins producing the guano patch visible in the aerial photograph. The secondary population of interest is all of the “local” penguins.
- A **sample** is taken from a population to provide estimate of population parameters. The sample of interest in this activity is the random sample of “local” penguins selected by the student. Estimates obtained from the “local” penguins can be used to infer information about the population of penguins in the aerial photograph.
- Different random samples will lead to different estimates of the mean guano produced. These will vary around the true mean of the population. The larger the sample, the smaller this variability will be.
- This activity can be extended to investigate the effects of non-random sampling (or the introduction of **bias**). The male penguins on the “Choose your penguin” activity sheet do more poo on average than the females. Consequently, taking non-random samples of just female penguins will result in a larger estimates of the number of penguins in the colony and vice versa. Another example of this could be sampling rockhopper penguins if the guano in the photograph comes from emperor penguins.

Risk assessment:

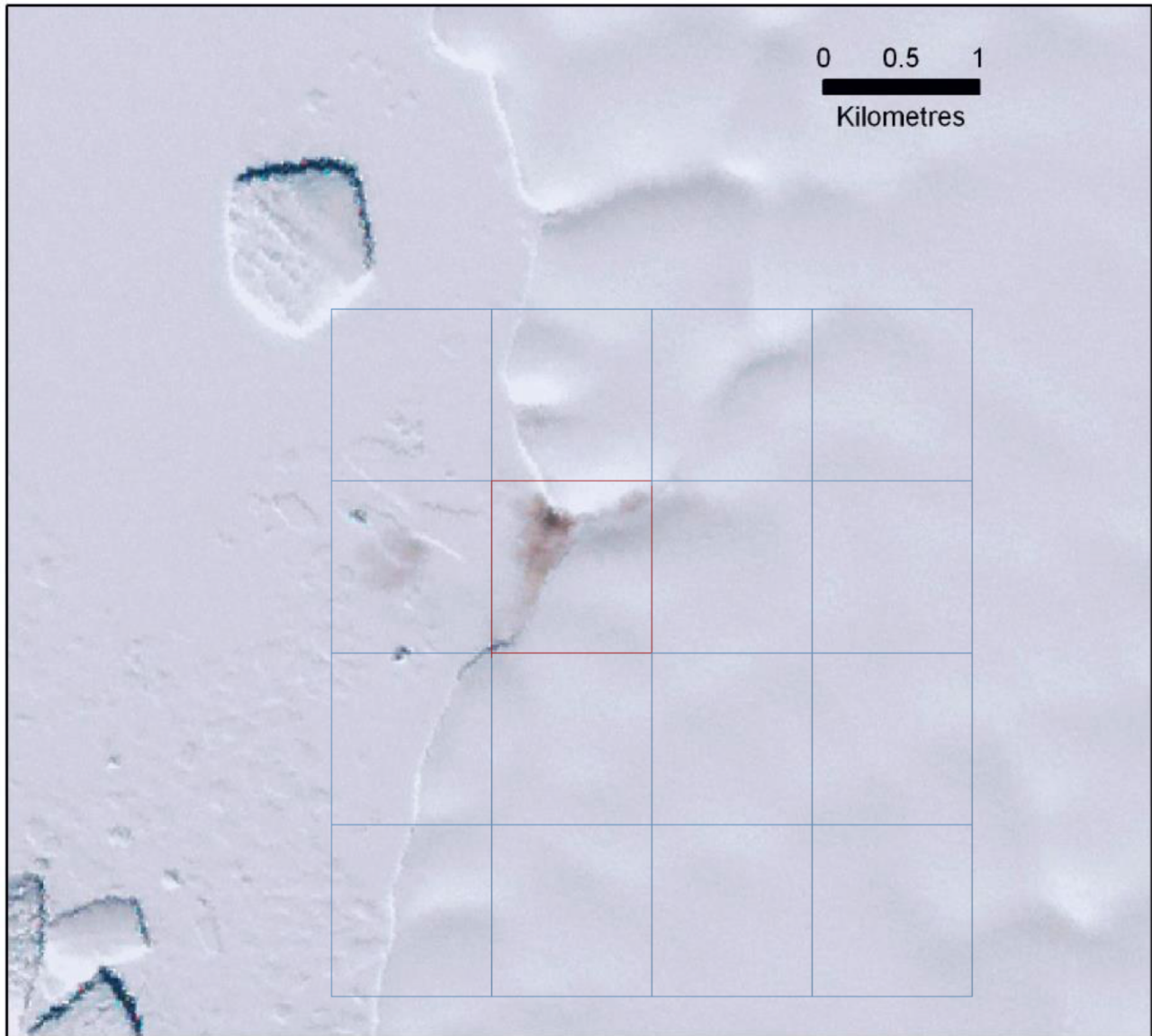
- There are no risks associated with this activity

Additional information and taking it further:

- StatTrek guide to samples and populations: <http://stattrek.com/sampling/populations-and-samples.aspx>
- Significance article describing an alternative way to estimate the penguin population: [The Emperor Strikes Back](#), Peter Fretwell. 12 June 2012 (Free access)

This activity was created by RSS fellows Dr Laura Bonnett and Dr Simon White, and adapted for MWS by RSS Edinburgh committee members Ailith Ewing, David Ewing and Tess Vernon.

Aerial photograph of penguin guano on snow



A medium-resolution Landsat image of an emperor penguin colony in West Antarctica.

Aerial photograph courtesy of DigitalGlobe

Choose your penguins!

Name: Roland
Poo area: 35m²



Name: Barry
Poo area: 30m²



Name: Serena
Poo area: 28m²



Name: Claire
Poo area: 25m²



Name: LeBron
Poo area: 42m²



Name: Winston
Poo area: 38m²



Name: Elsa
Poo area: 23m²



Name: Rupert
Poo area: 36m²



Name: Luna
Poo area: 33m²

