Columns / Fields available in count data download	
Column / Field name	Description / Definition
Site name	Full site name as defined by the ANTARCTIC SITE COMPENDIUM
Site id	Four letter site codes defined by Lynch for identification purposes across the database. This is the primary key across the tables • E.g. Acuna Island = ACUN
CCAMLR region	This is the numeric value for the CCAMLR sub-region that the site falls within. For a map of the regions go to the CCAMLR website
Longitude	The longitude of the centroid of the site in WGS 1984 (EPSG 4326) as calculated by way of a site polygon manually assigned by Lynch et al.
Latitude	The latitude of the centroid of the site in WGS 1984 (EPSG 4326) as calculated by way of a site polygon manually assigned by Lynch et al.
Common name	English common name of the species counted. Either: Adelie penguin, Chinstrap penguin, Emperor penguin or Gentoo penguin
Day	Day of the month that the count was performed if known. Some days may not be known due to the information not being avaiable in the publication or report
Month	Month of the year (numeric value: e.g. Jan = 01, Feb = 02) that the count was performed if known. Some months may not be known due to information being unavailable in the report or publication
Year	Gregorian calender year that the count was performed
Season starting	The season of the count. This does not always correspond to the year. Counts performed between Jan - June of year N are included in season = N – 1 • E.g. February 2010 is Season = 2009
Count	The actual count either derived or counted for the population. This is a whole number
Accuracy	This is a quality flag that is given to each count in the database. It is scaled from 1 to 5 with 1 being the highest quality and 5 being the lowest. Values of 1 are mostly associated with ground counts

	while values of 5 are associated with less accurate estimation techniques like satellite / VHR images.
Count type	This is the life stage counted and is either: • **Nests**: usually counted early in the season by counting individual adults clearly sitting on a nest • **Chicks**: Counted later in the season after the chicks have hatched and before they have clearly fledged (primarily individuals with downy feathers) • **Adults**: Counted at various times of the year, but could include molting birds, and non-breeders in a particular colony. May not be reflective of the number breeding
Vantage	The type of platform used for counting or estimating populations. This can be one of: • aerial: Estimates performed from aircraft • aerial photo: Counts performed using images taken from an aircraft • ground: Counts performed by individuals on the ground within a colony • ground photo: These are counts performed using a high-resolution photograph. Usually when colonies are too large, or time does not permit for a full count at that moment • landsat: estimates from the landsat satellite by estimating guano patch size and extrapolating • offshore vessel: Counts made by an observer from a stationary vessel platform near a colony. Performed when the colony is not accessible • None: These values occur when the information on how an estimate was performed was unavaialble in the report or publication. Efforts are underway to contact authors and individuals to clarify these data
Reference	This is the citation for the count wrapped with HTML code for formatting purposes. The formatting can be viewed when you click on "view source" in the count tables tab of your query
Notes	These are important notes for the particular count and may include information that will help determine how the accuracy or quality of the count was determined

Columns when downloading modeled data

Mean	The posterior samples from the Bayesian population model (n = 4500) are averaged across all sites in your query. E.g. Average(Sample1 Site1,...Sample1 SiteN), Average(Sample2 Site1,...Sample2SiteN),... Average(Sample4500 Site1,... Sample4500 SiteN). This gives 4500 "possible" values for the total population. We take the mean of those 4500 samples as the estimated population size
Lower CI	This is the lower bound of the 90th percentile of the Bayesian credible interval calculated from the 4500 samples we use to calculate the mean value (see above). This is calculated using the highest posterior density interval (used for Bayesian techniques) as opposed to standard confidence intervals
Upper CI	This is the upper bound of the 90th percentile of the Bayesian credible interval calculated from the 4500 samples we use to calculate the mean value (see above). This is calculated using the highest posterior density interval (used for Bayesian techniques) as opposed to standard confidence intervals

Metadata contact:

Grant Humphries: grwhumphries@blackbawks.net
Heather Lynch: heather.lynch@stonybrook.edu
Christian Che-Castaldo: checastaldo@icloud.com

Address: 106 Life Sciences. Department of Life Sciences. Stony Brook University. Stony Brook, NY, 99775

Polygons / Site Masks:

Projection: EPSG 3031
Creation: Polygons were created by hand in ArcGIS / ArcMap software. Each polygon represents the approximate boundary of the site. The site locations are determined by the centroid of these site masks.
Created by: Heather Lynch (see above contact)