

## ĐỀ THI GIỮA HỌC KỲ II – Thống kê máy tính

**Đề bài:** Khảo sát dữ liệu về 3 loài chim cánh cụt

Dữ liệu được cung cấp trong file ‘penguins-data.csv’ được khảo sát các số đo của 3 loài chim cánh cụt trên 3 đảo (Biscoe, Dream, Torgersen). Yêu cầu:

**Câu 1.** (1 điểm) Đọc dữ liệu và hiển thị 10 dòng dữ liệu đầu tiên

**Câu 2.** (5 điểm) Tìm hiểu thông tin về dữ liệu:

**1.1.** Số lượng chim cánh cụt của mỗi loài

Loài	Chinstrap	Gentoo	Adelie
Số lượng			

**1.2.** Số lượng mỗi loài chim cánh cụt trên mỗi đảo

	Chinstrap	Gentoo	Adelie
Biscoe			
Dream			
Torgersen			

**1.3.** Với mỗi loại chim cánh cụt, bạn hãy so sánh các đại lượng thống kê: giá trị trung bình, trung vị, độ lệch chuẩn trên các thuộc tính độ dài mỏ (bill\_length\_mm), độ sâu mỏ (bill\_depth\_mm), độ dài cánh (flipper) và trọng lượng (body\_mass).

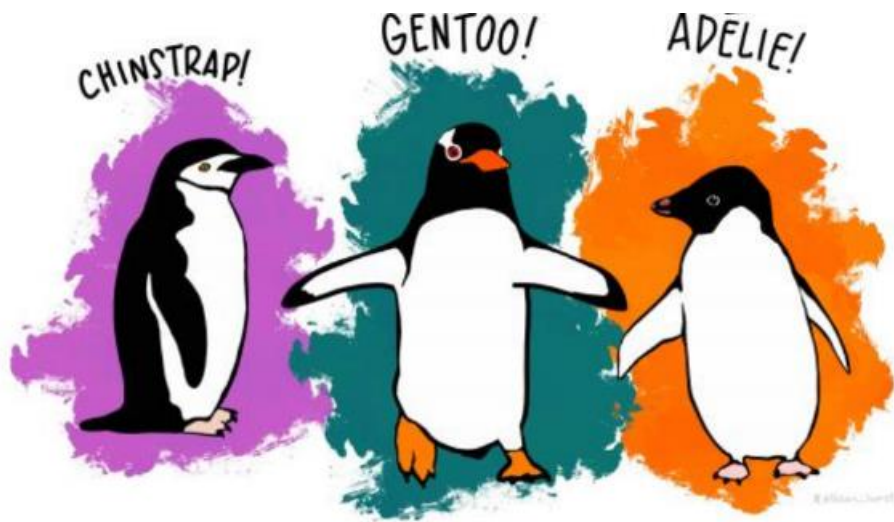
**1.4.** Dựa trên kết quả tính được, bạn có rút ra được nhận xét gì không?

**Câu 3.** (2 điểm) Vẽ đồ thị phù hợp thể hiện phân phối của ba loài chim cánh cụt, qua đó bạn hãy đưa ra bình luận của mình.

**Câu 4.** (2 điểm) Chọn loại đồ thị phù hợp để khảo sát mối quan hệ giữa độ dài cánh và trọng lượng của loài chim cánh cụt Gentoo:

**4.1.** Theo bạn hai đại lượng này có quan hệ với nhau hay không? Giải thích.

**4.2.** Ngoài độ dài cánh, theo bạn còn thuộc tính nào có khả năng ảnh hưởng đến trọng lượng của loài chim Gentoo?



## Body Part of Penguin

