

TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP
THÀNH PHỐ HỒ CHÍ MINH

KHOA CÔNG NGHỆ THÔNG TIN

ĐỀ THI GIỮA KỲ - ĐỀ 1

Môn thi: THỐNG KÊ MÁY TÍNH ỨNG DỤNG

Lớp/Lớp học phần: K16 (Kỳ 1, 2022-2023)

Thời gian làm bài: 75 phút

Họ và tên thí sinh; MSSV:; Số Máy:

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

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

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

a. Số lượng chim cánh cụt của mỗi loài.

Loài	Chinstrap	Gentoo	Adelie
Số lượng			

b. Số lượng mỗi loài chim cánh cụt trên mỗi đảo

	Chinstrap	Gentoo	Adelie
Biscoe			
Dream			
Torgersen			

c. 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).

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

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.

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:

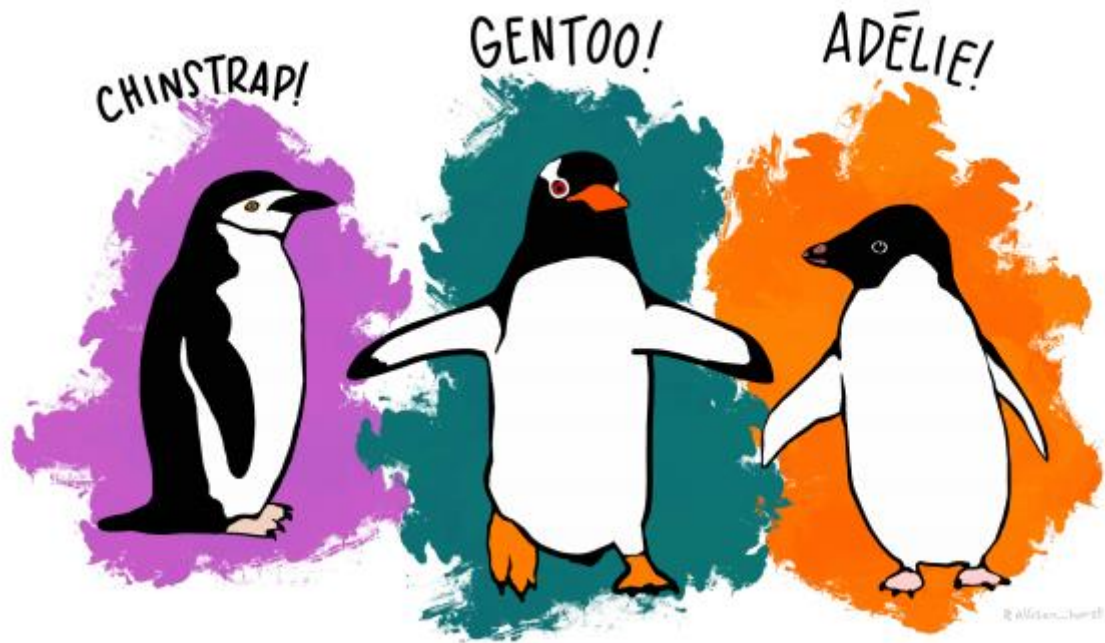
a. Theo bạn hai đại lượng này có quan hệ với nhau hay không? Giải thích.

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

Lưu ý:

- Bài thi lưu theo định dạng `<msv>_<hoten>.ipynb` và `<msv>_<hoten>.html`
- Phần đầu file bài làm ghi thông tin: mã sinh viên, họ tên, số máy.
- Các câu trả lời nhận xét, bình luận gõ trực tiếp vào file bài làm

----- Hết -----



Body Part of Penguin

