

Họ và tên thí sinh; MSSV:; Số Máy:

KHẢO SÁT DỮ LIỆU VỀ MỨC ĐỘ CÁC PHẢN HỒI CỦA HÀNH KHÁCH ĐỐI VỚI CÁC HÃNG MÁY BAY

Dữ liệu *TwitterUSAirlineSentiment.csv* bao gồm các cột.

airline_sentiment: Các mức độ phản hồi (positive, neutral, negative)

airline_sentiment_confidence: độ tin cậy của phản hồi

negativereason_confidence: độ tin cậy của lý do phủ định

airline: Hãng máy bay (Virgin America, United, Southwest, Delta, US Airways, American)

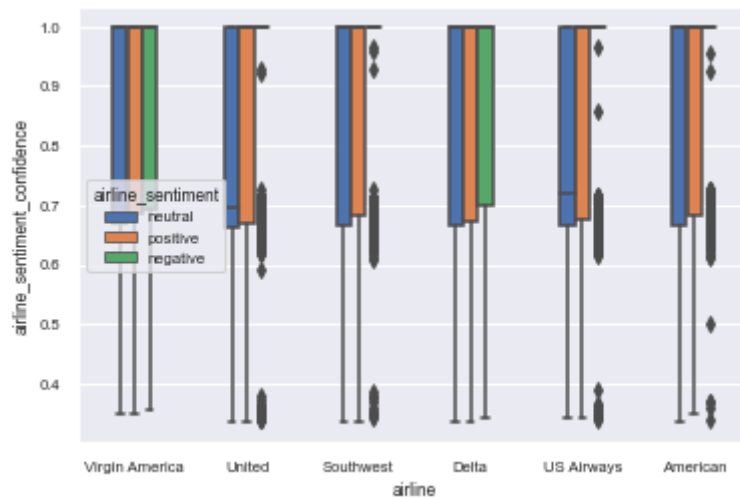
name: tên của hành khách phản hồi

text: phản hồi dạng chữ

user_timezone: múi giờ

Yêu cầu:

1. **(1 điểm)** Đọc dữ liệu và hiển thị 10 dòng dữ liệu đầu tiên
2. **(3 điểm)** Tìm hiểu thông tin về dữ liệu:
 - a. Số lượng phản hồi của mỗi hãng máy bay và số lượng phản hồi mỗi mức độ của từng hãng máy bay.
 - b. Với mỗi mức độ phản hồi của từng hãng máy bay, hãy xác định các tham số thống kê: giá trị trung bình, độ lệch chuẩn của các thuộc tính độ tin cậy của phản hồi (*airline_sentiment_confidence*), độ tin cậy của lý do phủ định (*negativereason_confidence*); và độ đo xu hướng tập trung (central tendency) của thuộc tính múi giờ (*user_timezone*).
3. **(3 điểm)** Đồ thị hóa dữ liệu:
 - a. Số lượng phản hồi của mỗi hãng máy bay bằng đồ thị tròn.
 - b. Bảng đồ thị hộp râu (box plot) biểu diễn trực quan về cách dữ liệu của thuộc tính độ tin cậy của phản hồi (*airline_sentiment_confidence*) theo hãng hàng không nằm rải rác trên mặt phẳng, so sánh giữa các mức độ phản hồi.



- c. Bảng đồ thị phù hợp biểu diễn sự tương quan bằng màu sắc giữa các thuộc tính dựa trên hệ số tương quan.
4. (3 điểm) Đường kính của một loại chi tiết do một máy sản xuất có phân phối chuẩn, kỳ vọng 20mm, độ lệch chuẩn 0,2mm). Lấy ngẫu nhiên 1 chi tiết máy. Tính xác suất để
- có đường kính nhỏ hơn 20,3mm
 - có đường kính trong khoảng 19,9mm đến 20,3mm,
 - có đường kính sai khác với kỳ vọng không quá 0,3mm.

Lưu ý:

- Bài thi lưu theo định dạng **<msv>_<hoten>_GK.ipynb** và **<msv>_<hoten>_GK.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