

## Bài thực hành số 2

Lưu ý: Chủ đề trong diễn đàn nhóm tạo ra dùng để nộp các bài tập, các bài báo cáo, các kết quả thảo luận,... trong suốt học kỳ. Các nhóm có thể tham khảo nội dung của các nhóm khác nhưng tuyệt đối không sao chép cho các bài kết quả của nhóm mình (đề cao tính trung thực).

Trong các buổi làm việc: yêu cầu làm việc theo nhóm, các nhóm có thể trao đổi, hỗ trợ nhau khi cần thiết. Nhóm có thể mang máy tính cá nhân cùng làm việc.

### **Câu 1:**

Tập tin Full\_Mark\_2020.csv lưu trữ điểm thi TN/THPT năm 2020. Số lượng thí sinh rất lớn (trên 1 triệu học sinh). Nhóm tải tập tin này từ LMS, sau đó tìm ứng dụng phù hợp để khảo sát nội dung của tập tin này.

Cho biết:

Có bao nhiêu học sinh dự thi?

Điểm tổng hợp khối KHTN và KHXH tính ra sao?

### **Câu 2:**

Hãy dùng phần mềm phù hợp để kiểm tra với báo cáo của Bộ GD&ĐT theo yêu cầu sau:

Số bài thi môn Toán là bao nhiêu?

Phổ điểm của môn Toán (0, 0.2, 0.4,...,9.8,10)

Tính điểm trung bình, điểm trung vị, mode, độ lệch chuẩn của môn Toán.

### **Câu 3:**

Tìm hiểu mã tỉnh thành, sau đó trích xuất ra 2 bộ dữ liệu: 1 bộ chứa dữ liệu của các thí sinh thuộc TpHCM và 1 bộ chứa dữ liệu thí sinh thuộc Hà Nội.

Với mỗi bộ dữ liệu nêu trên, thực hiện các yêu cầu sau:

Cho biết số lượng thí sinh dự thi.

Xây dựng phổ điểm môn Toán, vẽ biểu đồ, tính các giá trị trung bình, trung vị, độ lệch chuẩn và mode.

Xây dựng phổ điểm tổng cộng (Toán+Văn+Ngoại ngữ+ KHXH/KHTN).

Vẽ biểu đồ cho điểm tổng cộng và tính toán các giá trị trung bình, trung vị, độ lệch chuẩn và mode.

Gửi 2 tập tin thí sinh của TpHCM và Hà Nội vào thread của nhóm (trong buổi thực hành).

### **Câu 4:**

Cách xếp loại tốt nghiệp THPT như sau:

Điểm xếp loại = Tổng số điểm các môn thi / Tổng số môn thi.

Tất cả bài thi và các môn thi thành phần của bài thi tổ hợp đăng ký dự thi để xét công nhận tốt nghiệp đều đạt trên 1 điểm, và có điểm xếp loại từ 5 điểm trở lên được công nhận tốt nghiệp THPT.

+ Loại giỏi: điểm xếp loại từ 8,0 điểm trở lên; không bài thi nào có điểm dưới 7,0.

+ Loại khá: điểm xếp loại từ 6,5 điểm trở lên; không bài thi nào bị điểm dưới 6,0.

+ Loại trung bình: các trường hợp còn lại.

Thực hiện việc xếp loại tốt nghiệp (luôn cả trường hợp Không tốt nghiệp) cho thí sinh của TpHCM

Sau đó cho biết 1 số thông tin sau:

Số thí sinh đậu tương ứng với từng mức xếp loại, số thí sinh không TNTHPT.

Vẽ biểu đồ thể hiện số liệu của số liệu bên trên.

**Câu 5:** Phân thu hoạch

Phân công thực hiện, thực hiện ở nhà, thời hạn 1 tuần và gửi vào chủ đề trong diễn đàn.

- Báo cáo các kết quả của câu 1-2-3-4.
- Dựa vào các số liệu, biểu đồ đã báo cáo hãy đưa ra các nhận định, phán đoán liên quan đến thí sinh tại 2 điểm thi TpHCM và Hà Nội.
- Hướng dẫn cài đặt và làm quen ngôn ngữ lập trình Python. Minh họa việc dùng Python để thực hiện một số việc sau:
  1. Cho 1 dãy các số nguyên/thực: tính toán các giá trị trung bình, trung vị, mode, độ lệch chuẩn của dãy số .
  2. Cho 1 dãy các số nguyên/thực và 1 số nguyên K. Thực hiện rời rạc hóa dãy số nguyên này bằng thuật toán “Equal Width” với K khoảng. Ví dụ dãy số là 3, 5, 11, 12, 20, 24 và K=3. Như vậy  $(27-3)/3=7$ , do đó sẽ rời rạc hóa thành 3 khoảng [3,10),[10,17),[17,24]
  3. Cho 1 dãy các số nguyên/thực và 1 số nguyên K. Thực hiện rời rạc hóa dãy số nguyên này bằng thuật toán “Equal Frequency” với K khoảng. Ví dụ dãy số là 3, 5, 11, 12, 20, 24 và K=2. Như vậy mỗi khoảng sẽ có 2 phần tử → rời rạc hóa thành các khoảng [3,8),[8,16),[16,24]

Lưu ý: Nhóm có thể dùng bất kỳ ứng dụng để thực hiện báo cáo (Power Point, Word, Clip/Video,...). Phân thu hoạch này cần nộp lên diễn đàn trước buổi học lý thuyết kế tiếp: GV có thể mời 1 nhóm lên trình bày.