

THỐNG KÊ MÁY TÍNH & ỨNG DỤNG

Tập dữ liệu sử dụng trong bài kiểm tra này được cung cấp trong file *IT-salaries-in-EU-2018.csv* – Đây là kết quả khảo sát online từ thị trường nhân lực trong lĩnh vực CNTT tại Châu Âu vào năm 2018. Anh/chị được yêu cầu thực hiện các công việc dưới đây. Link:

1. **(1 điểm)** Đọc dữ liệu và hiển thị 5 dòng dữ liệu đầu tiên.
2. **(3 điểm)** Tìm hiểu thông tin về dữ liệu:
 - a. (1đ) Dữ liệu bao gồm những thuộc tính nào? Liệt kê theo nhóm các thuộc tính định tính, các thuộc tính định lượng và chọn *tham số thống kê mô tả* phù hợp để đo *độ tập trung dữ liệu* của mỗi nhóm.
 - b. (2đ) Có bao nhiêu loại công ty (*company type*) trong khảo sát? Liệt kê các loại công ty đó.
3. **(3 đ)** Cho biết tỷ lệ nam và nữ xuất hiện trong khảo sát? Chọn loại đồ thị phù hợp trực quan để làm nổi bật so sánh của 2 tỷ lệ này.
4. **(2.0 đ)** Để so sánh mức lương của nam và nữ trong ngành IT:
 - a. Tính đại lượng đặc trưng cho *độ tập trung* và *độ phân tán* về mức lương của nam và nữ. Bình luận kết quả tính được.
 - b. Hãy chọn đồ thị phù hợp để trực quan làm nổi bật sự so sánh mức lương nam và nữ. Đưa ra bình luận của anh/chị. Kết quả này có tương đồng với bình luận của anh/chị ở trên không?
5. **(1 điểm)** Những yếu tố (thuộc tính) nào ảnh hưởng đến mức lương của nhân lực IT? Dùng bằng chứng thống kê để giải thích cho nhận định đó của anh/chị.