

| **TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP**  **THÀNH PHỐ HỒ CHÍ MINH**    **KHOA CÔNG NGHỆ THÔNG TIN** |  |  |
| --- | --- | --- |

Họ tên: ……………………………Mã SV:…………………Số máy:………

Sinh viên nộp bài một file duy nhất ***HoTen\_MaSV.\****

Cho *dataset* điểm **THPT** năm 2022 của học sinh Việt Nam. Anh/chị hãy thực hiện các yêu cầu sau:

1. ***(3đ)*** Đọc file dữ liệu ***diem\_thi\_thpt\_2022.csv*** và thực hiện:
   1. **(0.75đ)** Hiển thị 10 dòng đầu tiên và 10 dòng cuối cùng của bảng dữ liệu.
   2. **(0.75đ)** Cho biết bảng dữ liệu có bao nhiêu dòng và bao nhiêu cột?
   3. **(0.75đ)** Cho biết thông tin về các cột dữ liệu và kiểu của từng cột trong bảng
   4. **(0.75đ)** Cho biết số học sinh bị thiếu điểm môn Ngữ Văn và môn Ngoại Ngữ
2. ***(2.25đ)*** Với cột điểm thi môn Toán, hãy tính các trị thống kê:
3. **(0.75đ)** Trung binh (*mean*),
4. **(0.75đ)** Độ lệch chuẩn (*standard deviation*)
5. **(0.75đ)** Và khoảng dữ liệu (*range*)
6. ***(1.5đ)*** Dựa vào dữ liệu cột **Khoi**, anh/chị hãy thực hiện:
   1. **(0.75đ)** Đếm số lượng học sinh thi khối "KHTN" và thi khối "KHXH"
   2. **(0.75đ)** Vẽ đồ thị Pie để so sánh mối tương quan giữa "KHTN" và "KHXH".
7. ***(1.75đ)***
8. **(0.75đ)** Liệt kê 20 học sinh có điểm Toán cao nhất
9. **(0.75đ)** Liệt kê 20 học sinh có điểm Toán thấp nhất
10. **(0.25đ)** Cho biết có bao nhiêu điểm 10 môn Toán.
11. ***(1.5đ)***
12. **(0.75đ)** Thống kê để hoàn thành bảng số liệu dưới đây
13. **(0.75đ)** Từ đó trực quan hóa dữ liệu **Bảng 1** bằng đồ thị **Bar**.

**Bảng 1**

| **Vùng** | **Số học sinh** |
| --- | --- |
| thành phố hà nội |  |
| thành phố hồ chí minh |  |
| thành phố đà nẵng |  |

------------------------- Hết ------------------------------

*Sinh viên* ***tuyệt đối*** *không trao đổi bài, các bài giống nhau bị điểm 0 Thực hành*