

TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HỒ CHÍ MINH KHOA CÔNG NGHỆ THÔNG TIN	ĐỀ THI GIỮA KỲ Môn thi: THỐNG KÊ MÁY TÍNH ỨNG DỤNG Lớp/Lớp học phần: (Kỳ 2, 2023-2024) Thời gian làm bài: 60 phút

Họ và tên thí sinh; MSSV:; Số Máy:

Đề bài: KHẢO SÁT DỮ LIỆU VỀ MUA BÁN BẤT ĐỘNG SẢN NĂM 2020 TẠI VIỆT NAM

Dữ liệu được cung cấp trong file *bds.csv* được thu thập năm 2020 tại Việt Nam.

1. **(1 điểm)** Đọc dữ liệu và hiển thị 10 dòng đầu tiên của dữ liệu
2. **(5 điểm)** Tìm hiểu thông tin về dữ liệu:
 - a. **(0.75 điểm)** Số lượng loại mỗi loại (type) bất động sản có trong dữ liệu

Loại BDS	Số lượng
nhà	
chung cư	
đất	
khác	

- b. **(2 điểm)**
 - i. **(0.5 điểm)** Số lượng bất động sản của mỗi loại (type) theo từng cách thức rao bán (poster_temp).

Cách rao bán	khác	môi giới
nhà		
chung cư		
đất		
khác		

- ii. **(1.5 điểm)** Vẽ biểu đồ tròn thể hiện phần trăm của từng loại bất động sản theo từng cách thức rao bán dựa vào bảng trên và đưa ra nhận xét (tối thiểu 2 nhận xét)
- c. **(1.0 điểm)**
 - i. Tạo bảng dữ liệu mới chỉ có 3 vùng (tp hồ chí minh, hà nội và đà nẵng) và cho biết số dòng của bảng dữ liệu mới

Đề thi gồm 2 trang - Đề thi được sử dụng tài liệu.

- ii. (0.75) Từ bảng dữ liệu trên hãy so sánh các đại lượng thống kê giá trị trung bình, trung vị, độ lệch chuẩn của giá BDS của mỗi vùng, vẽ biểu đồ cột và đưa ra nhận xét

d. (1.25 điểm)

- i. **(0.5 điểm)** Tạo bảng dữ liệu có diện tích (*acreage khác -1*) và cho biết số dòng của bảng dữ liệu mới.
- ii. **(1.0 điểm)** Vẽ biểu đồ thể hiện số BDS có diện tích trong mô tả (*acreage khác -1*) của mỗi loại poster và đưa ra nhận xét

3. (2 điểm)

- a. **(1 điểm)** Chuyển dữ liệu time thành kiểu datetime, vẽ biểu đồ thể hiện số lượng bất động sản được bán theo từng tháng trong năm 2020 và đưa ra nhận xét
- b. **(1 điểm)** Hãy tính xác suất để một bất động sản được bán trong mùa hè (từ tháng 6 đến tháng 8) năm 2020
4. **(2 điểm)** Hãy dựa vào bảng dữ liệu được cung cấp, tính xác suất để trong 50 bất động sản được bán có 25 bất động sản có giá trị lớn hơn 3 tỷ

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 -----