

**BÀI TẬP 1**  
**THỐNG KÊ MÁY TÍNH VÀ ỨNG DỤNG**  
**KÌ 1 2022-2023, ĐÀO TẠO TỪ XA**

Lưu ý: Nộp file Python Notebook. Ghi rõ thông tin sinh viên ở đầu file.

**Câu 1. (2.5 điểm)** Bảng dữ liệu sau ghi nhận chế độ dinh dưỡng của 2 nhóm đối tượng (tính theo ounce cho 1 người trong 1 tuần).

	Nhóm A	Nhóm B
White bread	26	12.5
Sugar	11.5	8
Potatoes	48.5	33.5
Fruit	13	25.5
Vegetables	21.5	30.5
Brown bread	5.0	8

- Vẽ biểu đồ quạt (pie chart) cho nhóm A.
- Vẽ biểu đồ thanh (bar chart) cho nhóm B.
- Vẽ biểu đồ so sánh 2 nhóm.
- Nhận xét.

**Câu 2. (2.5 điểm)** Bảng dữ liệu sau thống kê số lượng sinh viên có điểm thi trong các khoảng tương ứng của một kì thi (điểm theo thang 100).

Khoảng điểm	< 25	[25, 30)	[30, 40)	[40, 60)	[60, 75)	[75, 100]
Số lượng	2	20	56	140	122	95

- Lập bảng cho thấy tần số (frequency), tần suất (relative frequency), mật độ (density), tần số tích lũy (cumulative frequency), tần suất tích lũy (cumulative relative frequency).
- Vẽ biểu đồ histogram.
- Vẽ biểu đồ tần suất tích lũy.
- Nhận xét.

**Câu 3. (2.5 điểm)** Sau đây là số liệu điểm thi của một lớp (tính theo thang 20)

2, 3, 7, 7, 8, 8, 9, 10, 11, 11, 12, 12, 14, 14, 15, 17, 17, 18, 19, 20

- Tính trung bình (mean), trung vị (median), yếu vị (mode), phân vị (percentile) 10%, 90%.
- Tính phạm vi (range), khoảng tứ phân vị (IQR - interquartile range), phương sai (variance), độ lệch chuẩn (standard deviation).
- Nhận xét.

**Câu 4. (2.5 điểm)** Tập tin GDPpercapita.csv chứa dữ liệu GDP bình quân đầu người (tính theo đô la Mỹ) của các nước trong năm 2019 và 2020 (nguồn World Bank: <https://databank.worldbank.org/indicator/NY.GDP.PCAP.CD/1ff4a498/Popular-Indicators>).

- “Phân tích” GDP bình quân đầu người năm 2020.

b) “Phân tích so sánh” GDP bình quân đầu người năm 2020 so với năm 2019.

--- HẾT ---