

WRITTEN BY MR. HUYNH NAM

EMAIL: giangdayit@gmail.com

FB: www.facebook.com/mrhuynhnham

Website: www.huynhnham.com || www.giangdayit.com

LAB: TRỰC QUAN HÓA DỮ LIỆU

Nội dung: Thực hiện vẽ các biểu đồ dữ liệu với thư viện Matplotlib

Mục tiêu: Các loại biểu đồ

1. Scatter (phân tán)
2. Line (đường)
3. Biểu đồ tròn
4. Biểu đồ cột (thanh)
5. Gant

PHẦN 1: CÁC BIỂU ĐỒ THÔNG DỤNG

Hãy ôn tập vẽ các mẫu biểu đồ cơ bản trong slide

PHẦN 2: ỨNG DỤNG TRỰC QUAN HÓA DỮ LIỆU

Bài 1: Cho bảng dữ liệu

Tháng	Giá vàng (usd / ounce)	Giá dầu thô (usd/thùng)	Số người nhiễm nCovid-19
15/04/2020	1730	20,2	2.000.134
16/04/2020			
17/04/2020			
18/04/2020			
19/04/2020			
20/04/2020			
21/04/2020			
22/04/2020			
23/04/2020			
24/04/2020			

- a. Sử dụng internet để điền vào dữ liệu còn thiếu cho các ngày còn lại.
- b. Khai báo dữ liệu trên các cột thành 4 danh sách tương ứng: ds_date, ds_gold, ds_oil, ds_ncovid19
- c. Vẽ biểu đồ line biểu thị giá vàng theo thời gian.
- d. Vẽ biểu đồ line biểu thị giá vàng, giá dầu theo thời gian. Lưu ý: các đồ thị này cùng nằm trên một figure nhưng khác axes

WRITTEN BY MR. HUYNH NAM

EMAIL: giangdayit@gmail.com

FB: www.facebook.com/mrhuynhnham

Website: www.huynhnham.com || www.giangdayit.com

- e. Vẽ biểu đồ scatter biểu thị sự phân tán dữ liệu tương quan giữa [giá, vàng và giá dầu], [giá vàng, số người nhiễm covid] và [giá dầu, số người nhiễm covid]. Lưu ý: các đồ thị này cùng nằm trên một figure nhưng khác axes

Lưu ý: Yêu cầu đồ thị phải có ghi label, title cho các trục và hình vẽ, ngoài ra biểu đồ phải được thiết lập giá trị, co giãn kích thước phù hợp hiển thị dữ liệu rõ ràng.

Bài 2: Cho bảng dữ liệu nhập hàng tại hệ thống siêu thị Coopmart Tây Ninh như sau:

Mặt hàng	Số lượng (tấn)
Cà phê	186
Dầu ăn	66
Gạo	74
Yoghurt	224
Trứng	24
Sô cô la	201
Nước ép trái cây	151
Phô mai	121
Mặt hàng khác	20

- a. Vẽ biểu đồ cột thể hiện số lượng nhập hàng cho từng mặt hàng
- b. Vẽ biểu đồ tròn biểu thị tỉ lệ phần trăm nhập hàng. **Lưu ý:** Làm nổi bật phần trăm mặt hàng được nhập nhiều nhất.
- c. Vẽ biểu đồ đường thể hiện tỉ lệ phần trăm tích lũy của các mặt hàng theo thứ tự trên

Lưu ý: Yêu cầu đồ thị phải có ghi label, title cho các trục và hình vẽ, ngoài ra biểu đồ phải được thiết lập giá trị, co giãn kích thước phù hợp hiển thị dữ liệu rõ ràng.

WRITTEN BY MR. HUYNH NAM

EMAIL: giangdayit@gmail.com

FB: www.facebook.com/mrhuynhnam

Website: www.huynhnam.com || www.giangdayit.com

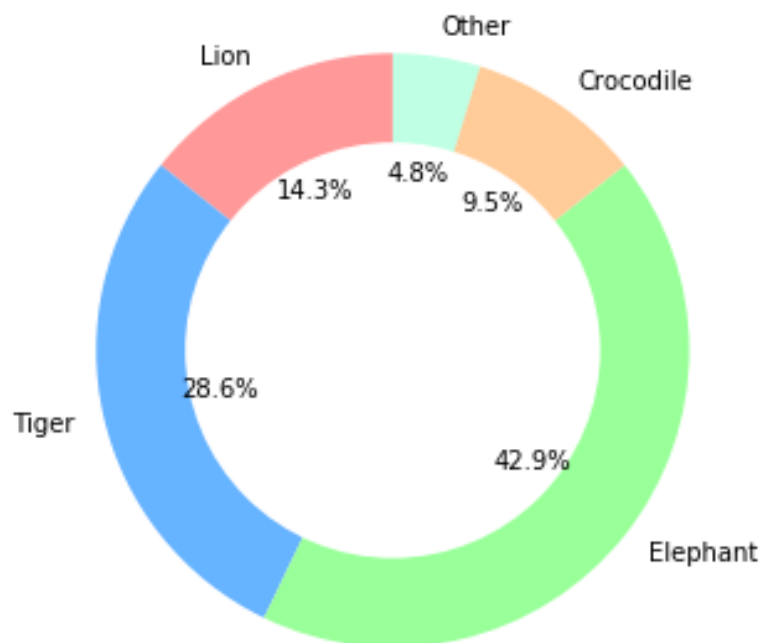
BÀI TẬP TƯ DUY: TRỰC QUAN HÓA DỮ LIỆU NÂNG CAO

1. Vẽ biểu đồ tròn nâng cao

Cho dữ liệu về số lượng các loài động vật quý hiếm trên thế giới. Hãy vẽ biểu đồ tỉ lệ phần trăm các loại động vật như hình

Animals	Amount (Thoundsand)
Lion	15
Tiger	30
Elephant	45
Crocodile	10

Kết quả



WRITTEN BY MR. HUYNH NAM

EMAIL: giangdayit@gmail.com

FB: www.facebook.com/mrhuynhnham

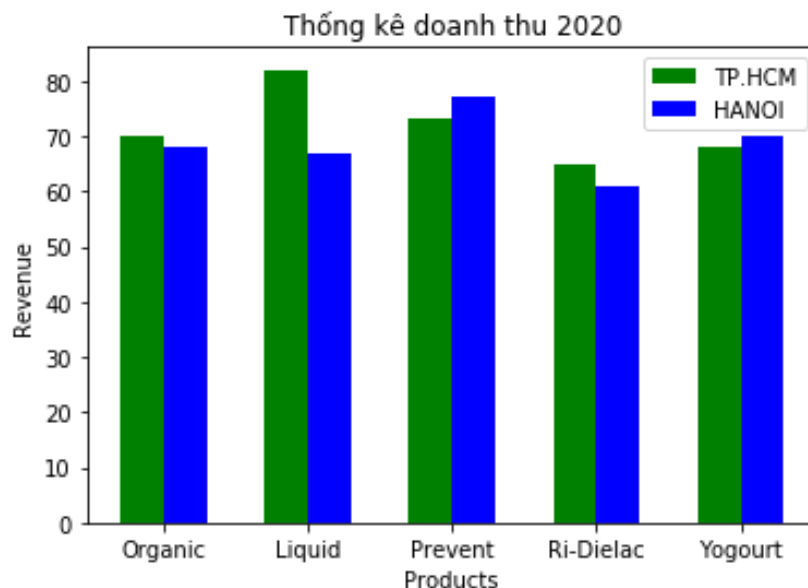
Website: www.huynhnham.com || www.giangdayit.com

2. Vẽ biểu đồ cột theo nhóm dữ liệu

Cho dữ liệu doanh thu (tỷ đồng) của các sản phẩm mũi nhọn của Vinamilk tại TP.HCM và Hà Nội như sau:

MẶT HÀNG	TP.HCM	HÀ NỘI
Organic	70	68
Liquid	82	67
Prevent	73	77
Ri-Dielac	65	61
Yogourt	68	70

Kết quả



3. Sử dụng biểu đồ 2-D để biểu diễn thông tin 3-D

Cho dữ liệu về cân nặng và chiều cao của học sinh trong lớp DHHTTT. Hãy vẽ biểu đồ phân tán thể hiện mối quan hệ giữa cân nặng và chiều cao. Biểu đồ cần thể hiện mức độ nghiêm trọng của sức khỏe (căn cứ vào chỉ số BMI) thông qua màu sắc và kích thước như sau: Dải màu thể hiện sức khỏe tốt đến nghiêm trọng tương ứng độ đậm nhạt của màu sắc. Mức độ gầy đến béo thể hiện tương ứng bán kính điểm vẽ từ nhỏ đến lớn.

WRITTEN BY MR. HUYNH NAM

EMAIL: giangdayit@gmail.com

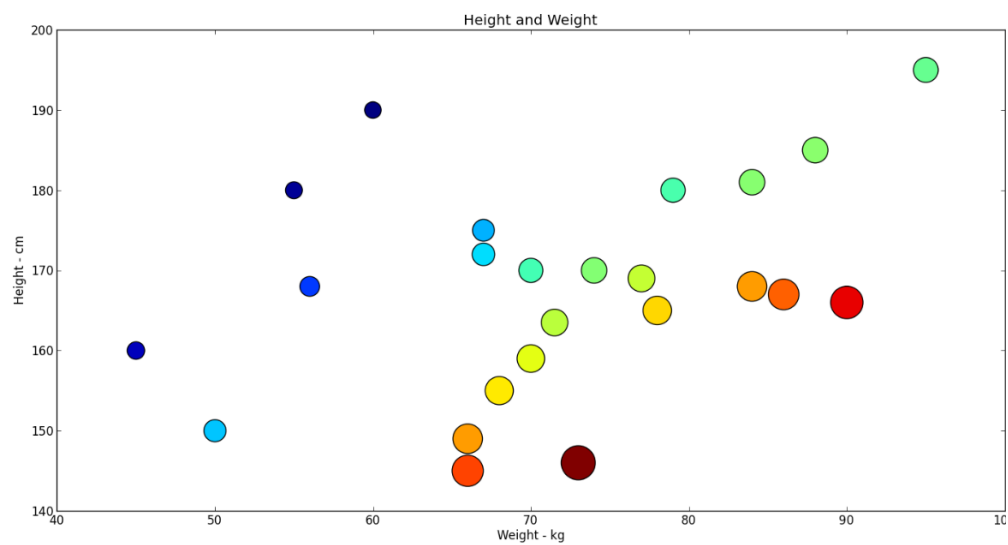
FB: www.facebook.com/mrhuynhnam

Website: www.huynhnam.com || www.giangdayit.com

Bảng dữ liệu

Weight (kg)	Height (cm)
55	180
86	167
74	170
66	149
78	165
68	155
79	180
90	166
73	146
70	159
88	185
66	145
84	168
67	172
84	181

Kết quả



WRITTEN BY MR. HUYNH NAM

EMAIL: giangdayit@gmail.com

FB: www.facebook.com/mrhuynhnham

Website: www.huynhnham.com || www.giangdayit.com

Bài 4: Viết code trực quan hóa biểu đồ Gannt biểu thị dữ liệu của các công việc được phân phối trong dự án sau (Xem tham khảo hình bên dưới)

Job	Start time (day)	Duration (days)
Job-2	10	50
Job-3	40	50
Job-1	60	30
Job-2	100	20
Job-1	110	10
Job-2	130	10
Job-1	150	10
Job-3	130	20

Hình tham khảo

