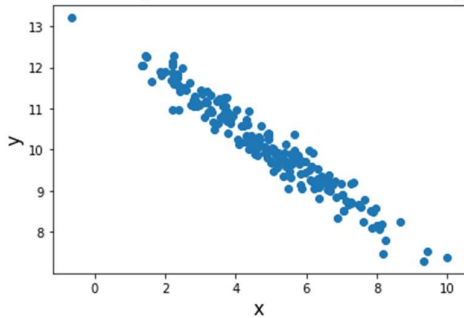


01	Khoa học dữ liệu là gì
	Là quá trình thu thập, lưu trữ, và quản lý dữ liệu
	Là việc sử dụng các mô hình thống kê để phân tích dữ liệu
	Là việc tạo ra các biểu đồ trực quan hoá dữ liệu để trình bày các quy luật trong dữ liệu
	Là lĩnh vực nghiên cứu liên ngành kết hợp toán thống kê, khoa học dữ liệu và một số kiến thức chuyên môn sâu của dữ liệu để tìm kiếm các quy luật của từ các tập dữ liệu
02	Đại lượng nào sau đây KHÔNG được coi là đặc trưng thống kê về cân nặng của đàn bò trong một điền trang
	Phương sai (variance)
	Tên người chủ điền trang
	Giá trị trung bình (mean)
	Giá trị trung vị (median)
03	Cho scatter plot biểu diễn hai biến: x, y như sau:
	
	Hỏi hệ số tương quan của x và y là:
	-1.00
	0.00
	-0.91
	1.00
04	Vào 2016 nhóm nghiên cứu Google DeepMind đã tạo ra AlphaGo đánh bại mọi kỳ thủ trong môn cờ nào?
	Cờ caro
	Cờ tướng
	Cờ vua
	Cờ vây
05	Thực hiện một bước tối ưu hoá của hàm: $y = x * x$, với $x_0 = 1.0$, $lr = 0.01$. Hỏi giá trị của x sau một bước tối ưu hoá là bao nhiêu
	0.98
	-0.98
	-0.64
	-0.97
06	Cho tập dữ liệu sau: $x = [-0.703, -0.344, -0.669, -0.636, -0.913]$ Giá trị trung bình của tập dữ liệu trên là?
	0.653

		-0.653
		0.365
		-0.365
07	Cho tập dữ liệu sau: $x = [-1.217, -2.059, -0.868, -0.341, -1.237]$ Giá trị độ lệch chuẩn (std) của tập dữ liệu trên là?	
		-0.651
		0.561
		0.615
		-0.615
08	Cho tập dữ liệu sau: $x = [-0.044, -0.314, -0.03, -0.01, -0.205]$ Giá trị TRUNG VỊ (median) của tập dữ liệu trên là?	
		-0.044
		-0.314
		-0.03
		-0.205
09	Nghiên cứu sự phụ thuộc của đại lượng X vào lượng y. Giả sử y phụ thuộc vào X như sau: $y_{\text{hat}} = a * X + b$. Sau khi huấn luyện mô hình, người ta thu được giá trị dự đoán như sau: $y_{\text{pred}} = [-0.519 -0.369 -1.129 -3.190 \ 0.299 \ 1.763 \ 4.586 \ 5.313 -5.038 -0.938]$. Biết giá trị thực tế của thu nhập như sau: $y = [-0.392 \ 0.267 -0.436 -2.957 -0.062 \ 1.234 \ 4.315 \ 5.610 -5.467 -1.333]$. Giá trị R2-score nào sau đây ĐÚNG?	
		0.980
		-0.980
		0.880
		0.780
10	Trong một dự án khoa học dữ liệu, người ta dự đoán khách hàng mua hàng và không mua hàng. Sau khi huấn luyện mô hình, người ta thu được giá trị dự đoán cho tập test như sau: $y_{\text{pred}} = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]$. Biết giá trị thực tế là: $y = [0 \ 1 \ 0 \ 1 \ 0 \ 0 \ 1 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]$. Giá trị nào sau đây ĐÚNG?	
		0.900
		0.950
		0.930
		0.990