Aula 03

Data Science

Estatística II

Variáveis

. Categórica

* Nominal
* Ordinal

Ordinal (masculino/feminino) → variável binária

a máquina compreende melhor -1 e 1

. Numéricas

* discretas → contagens (quantidade de pessoas na sala)
* contínuas

Regressor → valores contínuos

Classificadores

Probabilidade

1/50 milhões (ganhar na mega)

é um valor entre 0 e 1

Experimento Aleatório

jogar um dado de 6 lados → tenho chance iguais de tirar qualquer valores

game tearing

Espaço Amostral -> todas as possibilidades

Evento

Gráficos

Histograma (causa comum -> erro)

Boxplot → representa os quartis até os limites

quartil é dividir a população em 4 partes

não descartamos outliers

boxplot → só funciona se os dados estiverem normal

boxplot nao funciona na bimodal

Distribuição Normal

gaussiana bell curve

distribuição padrão

média - 0

desvio padrão - 1

Eventos aleatórios geram distribuição normal

Gallon Board - comprar

Usado em fraude

Nota de uma prova → colégio bernoulli

Normalização

divisão do mínimo pelo máximo

=(A6-2)/máximo

Normalização - Min Max

Espécie de uma planta

Normalização - Z Score

z = valores - média(valores) / desvio padrão

detector de IA do justdone verifica se o trabalho foi realizado por uma IA

humaniza textos

Cálculo da Probabilidade

Teste de Hipótese

Status Quo

Hipótese Nula = status quo

Hipótese Alternativa (provar)

alfa -> em quantos porcentos vou fazer minha afirmação (intervalo de erro)

tamanho da amostra (n)

média da amostra (x)

desvio padrão (conhecido)

Teste z -> distribuição normal

885-1000/0.01(raiz quadrada da amostra)

P-Valor

T-Student