

scRNA-seq: Cancer Microenvironment Classification

Features: 10
Amostras: 3000

Nome da Variável	Tipo	Descrição	Unidades	Dados faltantes
Cell_ID	Identificador	Código de identificação único para cada célula	Dado Categórico Nominal	Não foram observados valores ausentes na amostra
Cell_Type	Categórica	Classificação do tipo de célula (ex: T_Cell, Cancer, Fibroblast)	–	Não foram observados valores ausentes na amostra
Disease_Status	Categórica	Condição da amostra de origem da célula (ex: Tumor, Healthy_Control)	–	Não foram observados valores ausentes na amostra
Gene_E_Housekeeping	Numérica	Nível de expressão de um gene de manutenção ("housekeeping")	Adimensional	Não foram observados valores ausentes na amostra
Gene_A_Oncogene	Numérica	Nível de expressão de um oncogene (gene associado ao câncer)	Adimensional	Não foram observados valores ausentes na amostra
Gene_B_Immune	Numérica	Nível de expressão de um gene associado ao sistema imune	Adimensional	Não foram observados valores ausentes na amostra
Gene_C_Stromal	Numérica	Nível de expressão de um gene estromal (do tecido conjuntivo)	Adimensional	Não foram observados valores ausentes na amostra
Gene_D_Therapy	Numérica	Nível de expressão de um gene associado à resposta terapêutica	Adimensional	Não foram observados valores ausentes na amostra
Pathway_Score_Inflamm	Numérica	Pontuação calculada para a atividade de uma via de inflamação	Adimensional	Não foram observados valores ausentes na amostra
UMAP_1	Numérica	Primeira coordenada da	Adimensional	Não foram

		redução de dimensionalidade (UMAP)		observados valores ausentes na amostra
--	--	------------------------------------	--	--

Diabetes Dataset

Features: 9

Amostras: 768

Nome da Variável	Tipo	Descrição	Unidades	Dados faltantes
Pregnancies	Numérica	Número de vezes que a paciente ficou grávida	Contagem	“0” representa valor ausente
Glucose	Numérica	Concentração de glicose plasmática após 2 horas em um teste oral de tolerância à glicose	mg/dL	“0” representa valor ausente
BloodPressure	Numérica	Pressão arterial diastólica (mínima)	mm Hg	“0” representa valor ausente
SkinThickness	Numérica	Espessura da dobra cutânea do tríceps	mm	“0” representa valor ausente
Insulin	Numérica	Nível sérico de insulina após 2 horas	μU/mL	“0” representa valor ausente
BMI	Numérica	Índice de Massa Corporal (peso / altura ²)	kg/m ²	“0” representa valor ausente
DiabetesPedigreeFunction	Numérica	Função que estima a probabilidade de diabetes com base no histórico familiar	–	“0” representa valor ausente
Age	Numérica (inteira)	Idade da paciente	anos	“0” representa valor ausente
Outcome	Numérica	(binária) Indica se a paciente tem diabetes (1) ou não (0)		“0” representa valor ausente

Meu Município - Índice Brasileiro de Conectividade

Features: 13

Amostras: 22280

Nome da Variável	Tipo	Descrição	Unidades	Dados faltantes
Ano	Numérico	Ano de referência da coleta de dados	–	Não foram observados valores ausentes na amostra
Código Município	Numérico	Código oficial do município segundo o IBGE	–	Não foram observados valores ausentes na amostra
Município	Categórica	Nome do município	–	Não foram observados valores ausentes na amostra
UFC	Categórica	Sigla da Unidade da Federação (Estado) do município	–	Não foram observados valores ausentes na amostra
IBC	Numérico	Pontuação do Índice Brasileiro de Conectividade, que resume a qualidade da conexão no município	Adimensional	Não foram observados valores ausentes na amostra
Cobertura Pop. 4G5G	Numérico	Percentual da população do município com acesso às redes 4G ou 5G	Percentual	Não foram observados valores ausentes na amostra
Densidade SMP	Numérico	Densidade do Serviço Móvel Pessoal (telefonia móvel), geralmente medida em acessos por 100 habitantes	Acessos/100 habitantes	Não foram observados valores ausentes na amostra
HHI SMP	Numérico	Índice Herfindahl-Hirschman, que mede a concentração de mercado das operadoras de serviço móvel	Adimensional	Não foram observados valores ausentes na amostra
Densidade SCM	Numérico	Densidade do Serviço de Comunicação Multimídia (banda	Acessos/100 domicílios	Não foram observados valores ausentes

		larga fixa), geralmente medida em acessos por 100 domicílios		na amostra
HHI SCM	Numérico	Índice Herfindahl-Hirschman para o mercado de banda larga fixa	Adimensional	Não foram observados valores ausentes na amostra
Adensamento	Numérico	Mede a quantidade de estações (antenas) de telecomunicação em relação à população ou área.	Estações/10 mil habitantes	Não foram observados valores ausentes na amostra
Fibra	Numérico	Percentual dos acessos de banda larga fixa que utilizam tecnologia de fibra óptica	Percentual	Não foram observados valores ausentes na amostra
Cobertura área agricultável	Numérico	Percentual da área agricultável do município que possui cobertura de internet móvel	Percentual	“-” representa valor ausente