A

**A consulta externa:** Uma instrução SQL contendo uma subconsulta

**Abertura:** Aspecto da ética de dados que promove o acesso, uso e compartilhamento livre dos dados

**Agenda**: Lista de compromissos agendados

**Agregação de dados:** Processo de reunir dados de diferentes fontes e combiná-los em uma única coleção sintetizada

**Agregação:** Processo de coletar ou reunir diferentes partes em um todo

**Algoritmo:** Processo ou conjunto de regras seguidas para uma tarefa específica

**Amostra:** Em Data Analytics, um segmento de uma população que representa a população como um todo

**Amostragem aleatória:** Uma maneira de selecionar uma amostra de uma população para que cada tipo possível da amostra tenha a mesma chance de ser escolhido

**Amostragem não enviesada:** Quando a amostra da população medida representa a população como um todo

**Análise de dados:** A coleta, transformação e organização de dados para tirar conclusões, fazer previsões e orientar a tomada de decisão informada

**Análise de lacunas:** Método para examinar e avaliar o estado atual de um processo, a fim de identificar oportunidades para melhoria no futuro

**Analista de dados:** Alguém que coleta, transforma e organiza dados para tirar conclusões, fazer previsões e orientar a tomada de decisão informada

**Anonimização dos dados:** Processo de proteger dados privados ou sensíveis das pessoas, eliminando informações que possam identificá-los

**Anotação:** Texto que explica brevemente os dados ou ajuda a focar o público-alvo em um aspecto específico dos dados, em uma visualização

**Arquivo CSV (valores separados por vírgula):** Arquivo de texto delimitado que utiliza vírgulas para separar os valores

**Arquivo de áudio:** Gravação de áudio digitalizada, geralmente em formato MP3, AAC ou outro formato comprimido

**Arquivo de vídeo:** Coleção de imagens, arquivos de áudio e outros dados geralmente codificados em formato comprimido, como MP4, MV4, MOV, AVI, ou FLV

**Árvore de decisão:** Uma ferramenta que ajuda os analistas a tomarem decisões sobre pontos críticos de uma visualização

**Atribuição de alias:** A nomeação temporária de uma tabela ou coluna em uma consulta, a fim de facilitar a leitura e gravação

**Atributo:** Característica ou qualidade dos dados usada para rotular uma coluna em uma tabela

**Atributos pré-atencionais:** Os elementos de uma visualização de dados que um público-alvo reconhece automaticamente sem esforço consciente

**AVERAGE**: Função de planilha que retorna a média dos valores de um intervalo selecionado

**AVERAGEIF:** Uma função de planilha que retorna a média de todos os valores de células de determinado intervalo que atendem a uma condição especificada

B

**Banco de dados normalizado:** Banco de dados em que são armazenados apenas dados relacionados em cada tabela

**Banco de dados relacional:** Banco de dados que contém uma série de tabelas que podem ser conectadas para formar relacionamentos

**Banco de dados:** Coleção de dados armazenados em um sistema de computador

**Big data:** Conjuntos de dados grandes e completos, que envolvem longos períodos e permitem aos analistas de dados tratarem de problemas de negócios de longo alcance

**Blending de dados:** Um método do Tableau que combina dados a partir de múltiplas fontes de dados

**Bordas**: Linhas que podem ser adicionadas ao redor de duas ou mais células em uma planilha

C

**Cabeçalho:** Primeira linha em uma planilha, que rotula o tipo de dados em cada coluna

**Cálculo:** Um ramo da matemática que envolve o estudo das taxas de mudança e as mudanças entre valores que estão relacionados por uma função

**Campo calculado:** Novo campo em uma tabela dinâmica que realiza determinados cálculos com base nos valores de outros campos

**Campo:** Uma única informação de uma linha ou coluna de uma planilha; em uma tabela de dados, normalmente é uma coluna da tabela

**Canal:** Um aspecto visual ou variável que representa as características dos dados em uma visualização

**CASE:** Instrução SQL que retorna registros que atendem às condições ao incluir uma instrução if/then em uma consulta

**CAST:** Função SQL que converte dados de um tipo de dados para outro

**Causa raiz:** Razão pela qual um problema ocorre

**Causalidade:** Quando uma ação leva diretamente a um resultado, como uma relação causa-efeito

**Chave estrangeira:** Campo na tabela do banco de dados que é uma chave primária em outra tabela (Consulte chave primária)

**Chave primária:** Identificador em um banco de dados que faz referência a uma coluna na qual cada valor é exclusivo (Consulte chave estrangeira)

**Ciclo de vida dos dados:** Sequência de etapas pelas quais os dados passam, e que incluem planejamento, coleta, gestão, análise, arquivamento e destruição

**Ciência de dados:** Área de pesquisa que utiliza dados brutos para criar novas maneiras de modelar e entender o desconhecido

**Classificação:** Processo de organizar os dados em uma ordem significativa para torná-los mais fáceis de entender, analisar e visualizar

**Cluster:** Uma coleção de pontos de dados em uma visualização de dados com valores similares

**COALESCE:** Função SQL que retorna valores não nulos em uma lista

**Combo de gráficos:** Uma visualização de dados que combina mais de um tipo de visualização

**Compatibilidade:** Como dois ou mais conjuntos de dados são capazes de trabalhar em conjunto

**Composição dos dados:** Processo de combinar partes individuais em uma visualização e então mostrar essas partes juntas como um todo

**Comprimento de campo:** Ferramenta para determinar quantos caracteres podem ser digitados em um campo da planilha

**Comprimento:** Número de caracteres em uma string de texto

**CONCAT:** Função SQL que adiciona strings para criar strings de texto novas que podem ser usadas como chaves exclusivas

**CONCATENATE:** Função de planilha que une duas ou mais strings de texto

**Conjunto de dados:** Coleção de dados que pode ser manipulada ou analisada como uma unidade única

**Consentimento:** Aspecto da ética de dados que presume o direito que um indivíduo tem de saber como e por que seus dados pessoais serão usados antes de concordar em fornecê-los

**Consistência:** Grau em que os dados são repetíveis de diferentes pontos de entrada ou coleta

**Construção de relações profissionais:** Construindo relacionamentos conhecendo pessoas tanto pessoalmente quanto online (Consulte Networking)

**Consulta interna:** Uma subconsulta do SQL que está dentro de outra instrução SQL

**Consulta:** Solicitação de dados ou informações de um banco de dados

**CONTAGEM**: Uma função de planilha que conta o número de células dentro de uma faixa que atende a uma condição específica

**Contexto:** Condição em que algo existe ou acontece

**Controle de acesso:** Funcionalidades como proteção por senha, permissões do usuário e criptografia, usadas para proteger uma planilha

**Convenções de nomenclatura:** Diretrizes consistentes que descrevem o conteúdo, a data de criação e a versão de um arquivo em seu nome

**CONVERT:** Uma função SQL que muda a unidade de medida de um valor em dados

**Cookie:** Arquivo pequeno armazenado em um computador, contendo informações sobre seus usuários

**Correlação:** A medida do grau em que duas variáveis mudam em relação uma à outra

**COUNT DISTINCT:** Função do SQL que retorna somente os valores distintos de determinado intervalo

**COUNTA:** Uma função de planilha que conta o número total de valores dentro de uma faixa que atende a critérios especificados

**COUNTIF:** Função de planilha que retorna o número de células em um intervalo que corresponde a um valor especificado

**CRIE UMA TABELA:** Uma cláusula do SQL que adiciona uma tabela temporária a um banco de dados que pode ser usado por várias pessoas

D

**Dados abertos:** Dados que estão disponíveis para o público

**Dados amplos:** Conjunto de dados em que cada assunto possui uma única linha com colunas múltiplas a fim de conter os valores de diversos atributos do assunto

**Dados booleanos:** Dados com apenas dois valores possíveis, geralmente verdadeiro ou falso

**Dados contínuos:** Dados que são medidos e que podem ter quase qualquer valor numérico

**Dados de terceiros:** Dados fornecidos de fontes externas que não os coletaram diretamente

**Dados desatualizados:** Quaisquer dados que tenham sido substituídos por informações mais recentes e mais precisas

**Dados dinâmicos:** Dados que são atualizados automaticamente

**Dados discretos:** Dados contados e que têm um número limitado de valores

**Dados duplicados:** Qualquer registro que, inadvertidamente, compartilhe dados com outro registro

**Dados estáticos:** Dados que não mudam depois de registrados

**Dados estruturados:** Dados organizados em determinado formato, como em linhas e colunas

**Dados externos:** Dados que existem e são gerados fora de uma organização

**Dados incompletos:** Dados onde estão faltando campos importantes

**Dados inconsistentes:** Dados que utilizam formatos diferentes para representar a mesma coisa

**Dados incorretos/imprecisos:** Dados completos mas imprecisos

**Dados internos:** Dados que residem nos próprios sistemas da empresa.

**Dados limpos:** Dados completos, corretos e relevantes para o problema a ser resolvido

**Dados longos:** Conjunto de dados no qual cada linha é um ponto temporal por assunto, de modo que cada assunto possui dados em múltiplas linhas

**Dados não estruturados:** Dados que não são organizados de modo a serem facilmente identificados

**Dados nominais:** Tipo de dados qualitativos categorizados sem uma ordem definida

**Dados ordinais:** Dados qualitativos com uma ordem ou escala definida

**Dados pequenos** Pontos de dados pequenos e específicos, normalmente envolvendo um curto período de tempo, que são úteis para tomar decisões no dia a dia

**Dados primários:** Dados coletados por um indivíduo ou grupo usando seus próprios recursos

**Dados qualitativos:** Medida subjetiva e explicativa de uma qualidade ou característica

**Dados quantitativos:** Medida específica e objetiva, como um número, quantidade ou intervalo

**Dados secundários:** Dados coletados por um grupo diretamente de seu público-alvo, e vendidos posteriormente

**Dados sujos:** Dados incompletos, incorretos, ou irrelevantes para o problema a ser resolvido

**Dados tipo string:** Sequência de caracteres e pontuações que contêm informações textuais (também denominados dados tipo texto)

**Dados tipo texto:** Sequência de caracteres e pontuações que contêm informações textuais (também denominados dados tipo string)

**Dados:** Uma coleção de fatos

**Data Analytics:** Ciência dos dados

**DATEDIF:** Função de planilha que calcula o número de dias, meses ou anos entre duas datas

**Delimitador:** Caractere que indica o início ou o fim de um item de dados

**Design de dados:** O modo como a informação é organizada

**Design thinking:** Um processo usado para resolver problemas complexos de uma forma centrada no usuário

**Diagrama de dispersão:** Uma visualização de dados que representa os relacionamentos entre diferentes variáveis com pontos isolados de dados sem uma linha que os conecte

**Diagrama em caixa:** Uma visualização de dados que mostra a distribuição dos valores ao longo de um eixo x

**DISTINCT:** Palavra-chave que é adicionada a uma instrução SELECT SQL para extrair apenas entradas não duplicadas

**Domínio do problema:** Área de análise que engloba todas as atividades que afetam ou são afetadas por um problema

**DROP TABLE:** Uma cláusula SQL que remove uma tabela temporária de um banco de dados

E

**Ecossistema de dados:** Os vários elementos que interagem entre si para produzir, gerenciar, armazenar, organizar, analisar e compartilhar dados

**Eixo X:** A linha horizontal de um gráfico, geralmente posicionada na base, que frequentemente é usada para representar escalas de tempo e categorias discretas

**Eixo Y:** A linha vertical de um gráfico, geralmente posicionada à esquerda, que frequentemente é usada para representar frequências e outras variáveis numéricas

**Elemento de dados:** Uma informação dentro de um conjunto de dados

**Ênfase:** O princípio do design de organizar os elementos visuais para focar a atenção do público-alvo nas informações importantes em uma visualização de dados

**Engajamento:** Capturar e reter o interesse e a atenção de alguém durante uma apresentação de dados

**Engenheiro de dados:** Profissional que transforma os dados em um formato útil para análise e lhe dá uma infraestrutura confiável

**Equação:** Cálculo que envolve adição, subtração, multiplicação ou divisão (também chamada de expressão matemática)

**Equilíbrio:** O princípio do design de criar apelo estético e clareza em uma visualização de dados, distribuindo uniformemente os elementos visuais

**Escopo de trabalho (SOW):** Delineamento acordado das tarefas a serem realizadas durante um projeto

**Especialista em armazenamento de dados:** Profissional que desenvolve processos e procedimentos para efetivamente armazenar e organizar dados

**Esquema:** Uma maneira de descrever de que forma algo, como os dados, é organizado

**Estatísticas:** O estudo de como coletar, analisar, resumir, e apresentar dados

**Estratégia de dados:** Gestão de pessoas, processos e ferramentas usadas na análise de dados

**Estrutura:** O contexto de uma apresentação precisa criar conexões lógicas que se liguem à tarefa de negócios e à métrica

**Ética de dados:** Padrões bem fundamentados de certo e errado, que ditam como os dados são coletados, compartilhados e usados

**Ética:** Padrões fundamentados de certo e errado, que prescrevem o que os humanos devem fazer, geralmente em termos de direitos, obrigações, benefícios sociais, imparcialidade ou virtudes específicas

**Expressão matemática:** Cálculo que envolve adição, subtração, multiplicação ou divisão (também chamado de equação)

**Expressão regular (RegEx):** Regra que diz que os valores em uma tabela devem corresponder a um padrão prescrito.

F

**Filtragem:** Processo de exibir apenas os dados que atendem a um critério especificado, ocultando o restante

**Filtro de painel:** Uma ferramenta para mostrar apenas os dados que satisfazem um critério específico enquanto esconde o resto

**Flutuante:** Número que contém um decimal

**Fonte de dados boas:** Fonte de dados que é confiável, original, abrangente, atual e citada (ROCCC, na sigla em inglês)

**Fonte de dados inválidos:** Fonte de dados que não é confiável, original, abrangente, atual e citada (ROCCC, na sigla em inglês)

**Forma visual:** A aparência de uma visualização de dados que dá a ela estrutura e apelo estético

**Formatação condicional:** Ferramenta de planilha que muda a forma como as células aparecem quando os valores atendem a condições específicas

**Fórmula:** Conjunto de instruções usadas para realizar um cálculo usando os dados de uma planilha

**Foto digital:** Uma imagem eletrônica ou computadorizada geralmente em formato BMP ou JPG

**FROM:** Seção de uma consulta que indica de qual(is) tabela(s) para extrair os dados

**Função matemática:** Função usada como parte de uma fórmula matemática

**Função:** Comando predefinido que executa automaticamente um processo ou tarefa específica usando os dados de uma planilha

**Fusão de dados:** Processo de combinar dois ou mais conjuntos de dados em um único conjunto de dados

**Fusão:** Acordo que une duas organizações em uma única nova

G

**Geolocalização:** Localização geográfica de uma pessoa ou de um dispositivo, por meio da informação digital

**Governança dos dados:** Processo que garante a gestão formal dos recursos de dados de uma empresa

**Gráfico bolha concentrado:** Uma visualização de dados que mostra dados em círculos agrupados

**Gráfico de área:** Uma visualização de dados que usa pontos isolados dos dados para mudar uma variável conectada por meio de uma linha contínua com uma área preenchida por baixo

**Gráfico de barras:** Uma visualização de dados que usa o tamanho para contrastar e comparar dois ou mais valores

**Gráfico de bolha:** Uma visualização de dados que mostra pontos isolados dos dados como bolhas, comparando valores numéricos pelo seu tamanho relativo

**Gráfico de colunas:** Uma visualização de dados que usa pontos de dados individuais de uma variável dinâmica, representados como colunas verticais

**Gráfico de distribuição:** Uma visualização de dados que mostra a frequência de vários resultados em uma amostra

**Gráfico de Gantt:** Uma visualização de dados que mostra a duração de eventos ou atividades em uma linha do tempo

**Gráfico de linhas:** Uma visualização de dados que usa uma ou mais linhas para mostrar mudanças ou alterações nos dados ao longo do tempo

**Gráfico de marcador:** Uma visualização de dados que mostra os dados como um gráfico de barras horizontais movendo-se até um valor desejado

**Gráfico de medidor:** Uma visualização de dados que mostra um único resultado dentro um intervalo progressivo de valores

**Gráfico de pizza:** Uma visualização de dados que usa segmentos de um círculo para representar as proporções de cada categoria de dados comparada com o todo

**Gráfico de rosca:** Uma visualização de dados em que as proporções 2D ou 3D (segmentos) de um anel ou formato O representam valores de dados somando um todo ou 100%

**Gráfico dinâmico:** Gráfico criado a partir dos campos de uma tabela dinâmica

**Gráfico:** Uma representação gráfica dos dados de uma planilha

**GROUP BY:** Cláusula de SQL que agrupa linhas que têm os mesmos valores de uma tabela em linhas de resumo

H

**Habilidades analíticas:** Qualidades e características associadas ao uso de fatos para solucionar problemas

**Habilidades comportamentais:** Características e comportamentos não técnicos relacionados à maneira como as pessoas trabalham

**Habilidades transferíveis:** Habilidade e qualidades que podem ser transferidas de um trabalho a outro ou de um setor a outro

**HAVING** Uma cláusula no SQL que adiciona um filtro a uma consulta em vez da tabela subjacente que só pode ser usada com funções de agregação

**Hipótese:** Uma teoria que se pode tentar provar ou refutar com dados

**Histograma:** Uma visualização de dados que mostra com que frequência os valores de dados se enquadram em determinados intervalos

**História:** A narrativa de uma apresentação de dados que a torna interessante e significativa

I

**Imparcialidade:** Qualidade da análise de dados que não cria ou reforça um viés

**INNER JOIN:** Função de SQL que retorna registros com valores correspondentes em ambas as tabelas

**Integridade dos dados:** A precisão, integridade, consistência e confiabilidade dos dados ao longo de seu ciclo de vida

**Integridade:** Grau em que os dados contêm todos os componentes ou medidas desejados

**Interoperabilidade dos dados:** Capacidade de integrar dados de múltiplas fontes, que é um fator importante que leva ao uso bem-sucedido de dados abertos entre empresas e governos

**Intervalo de confiança:**  Um intervalo de valores que transmite a probabilidade de uma estimativa estatística refletir a população

**Intervalo de dados:** Valores numéricos que se situam entre valores máximos e mínimos predefinidos

**Intervalo:** Conjunto de duas ou mais células em uma planilha

J

**JOIN:** Função de SQL que é usada para combinar linhas de duas ou mais tabelas com base em uma coluna relacionada.

K

L

**Label:** Texto em uma visualização que identifica um valor ou descreve uma escala

**LEFT JOIN:** Função do SQL que retornará todos os registros da tabela à esquerda e somente os registros correspondentes da tabela à direita

**LEFT:** Função que retorna um determinado número de caracteres do lado esquerdo de uma string de texto

**Legenda:** Uma ferramenta que identifica o significado de vários elementos em uma visualização de dados

**LEN:** Função que retorna o comprimento de uma string de texto, contando o número de caracteres que ela contém

**LIMIT:** Cláusula de SQL que especifica o número máximo de registros retornados em uma consulta

**Linguagem de consulta estruturada:** Linguagem de programação computacional usada para se comunicar com um banco de dados

**Linguagem de consulta:** Linguagem de programação computacional usada para se comunicar com um banco de dados

**Localizar e substituir:** Ferramenta que localiza um determinado termo de pesquisa e o substitui por outro termo

**Log de mudanças:** Arquivo que contém uma lista em ordem cronológica de todas as modificações feitas no projeto

M

**Manipulação de dados:** Processo de mudança de dados para torná-los mais organizados e mais fáceis de ler

**Mapa de densidade:** Uma visualização de dados que representa concentrações, com cores representando o número ou frequência de pontos de dados em uma dada área em um mapa

**Mapa de símbolos:** Uma visualização de dados que mostra uma marca sobre uma longitude e uma latitude determinadas

**Mapa preenchido:** Uma visualização de dados que colore áreas em um mapa com base em medidas ou dimensões

**Mapa:** Uma visualização de dados que organiza os dados geograficamente

**Mapas de concentração:** Uma visualização de dados que usa contrastes de cor para comparar categorias em uma base de dados

**Mapeamento dos dados:** Processo de correspondência de campos de uma fonte de dados para outra

**Marca:** Um objeto visual em uma visualização de dados, tal como uma forma, um ponto, ou linha

**Margem de erro:** Quantidade máxima que se espera que os resultados da amostra sejam diferentes dos da população real

**Margem de lucro:** Porcentagem que indica quantos centavos de lucro foram gerados para cada dólar de venda

**MATCH:** Função de planilha usada para localizar a posição de determinado valor de pesquisa

**Matriz:** Coleção de valores em células da planilha

**MAX:** Uma função que retorna o maior valor numérico a partir de uma gama de células

**MAXIFS:** Uma função de planilha que retorna o valor máximo de um determinado intervalo que atende a uma condição especificada

**Mentalidade técnica:** Capacidade de dividir itens em etapas ou partes menores e trabalhar com elas de maneira ordenada e lógica

**Mentor:** Alguém que partilha seu conhecimento, suas habilidades e sua experiência para ajudar alguém a crescer profissionalmente e pessoalmente

**Meta métrica:** Meta mensurável definida por uma empresa e é avaliada com o uso de métricas

**Metadados administrativos:** Metadados que indicam a fonte técnica de um ativo digital

**Metadados descritivos:** Metadados que descrevem uma parte dos dados e que podem ser usados para identificá-lo posteriormente

**Metadados estruturais:** Metadados que indicam como uma parte dos dados é organizada e se faz parte de uma ou mais de uma compilação de dados

**Metadados:** Dados a respeito de dados

**Metodologia SMART:** Ferramenta para determinar a eficácia de uma pergunta, verificando se ela é específica, mensurável, orientada para a ação, relevante e limitada pelo tempo

**Métrica de negócios:** Um padrão de medição utilizado para resolver uma tarefa de negócios

**Métrica:** Tipo de dados quantificável e isolado que pode ser usado para medição

**MID:** Função que retorna um segmento do meio de uma cadeia de texto

**Mídias sociais:** Websites e aplicativos por meio dos quais os usuários criam e compartilham conteúdo ou participam de redes sociais

**MIN**: Função de planilha que retorna o menor valor numérico de um intervalo de células

**MINIFS:** Uma função de planilha que retorna o valor mínimo de um determinado intervalo que atende a uma condição especificada

**Modelo de dados:** Ferramenta para organizar os elementos dos dados e como eles se relacionam entre si

**Modelos mentais:** A linha de pensamento de um(a) analista de dados e a forma como ele(a) aborda um problema

**Módulo:** Operador (%) que retorna o resto quando um número é dividido por outro

**Moeda de troca:** O aspecto da ética de dados que presume os indivíduos devem estar cientes das transações financeiras resultantes do uso de seus dados pessoais e da escala dessas transações

**Movimento:** O princípio do design de arranjar os elementos visuais para guiar os olhos do público-alvo de uma parte da visualização de dados para outra

N

**Narrativa:** (Consulte a história)

**Networking:** Construção de relacionamentos por meio de encontros presenciais e online com outras pessoas

**Nível de confiança:** A probabilidade de que o tamanho de uma amostra reflita com precisão a maior população

**Notebook:** Ambiente interativo e editável de programação para criação de relatórios de dados e para exibição das habilidades dos dados

**Nulo:** Indicação de que um valor não existe em um conjunto de dados

**Nuvem:** Lugar para armazenar os dados online, em vez de no disco rígido de um computador

O

**O Método McCandless:** Um método para apresentar visualizações de dados que passa de informações gerais para informações específicas

**Obrigatório:** Valor de dados que não pode ser deixado em branco ou vazio

**Observação:** Atributos que descrevem um dado contido em uma linha de uma tabela

**Operador:** Símbolo que nomeia a operação ou o cálculo a ser realizado

**Ordem das operações:** Uso de parênteses para agrupar os valores das planilhas para esclarecer a ordem em que as operações devem ser realizadas

**ORDER BY:** Cláusula do SQL que classifica os resultados retornados em uma consulta

**Organização Mundial da Saúde:** Organização cujo principal papel é dirigir e coordenar a saúde internacional dentro do sistema da Organização das Nações Unidas

**OUTER JOIN:** Função de SQL que combina as funções RIGHT e LEFT JOIN para retornar todos os registros correspondentes em ambas as tabelas.

P

**Padrão:** O princípio do design de usar elementos visuais similares para demonstrar tendências e relacionamentos em uma visualização de dados

**Painel:** Ferramenta que monitora dados recebidos em tempo real

**Paleta divergente de cores:** Um tema colorido que mostra dois intervalos de valores dos dados usando duas matizes diferentes, com a intensidade da cor representando a magnitude dos valores

**Partes interessadas:** Pessoas que investem tempo e recursos em um projeto e estão interessadas em seu resultado

**Patrocinador:** Profissional comprometido a dar suporte ao avanço da carreira de outro profissional

**Pensamento analítico:** Processo de identificar e definir um problema e, em seguida, resolvê-lo usando dados de maneira organizada seguindo um passo a passo

**Pensamento estruturado:** Processo de reconhecer o problema ou situação atual, organizar as informações disponíveis, revelar lacunas e oportunidades e identificar as opções

**Pergunta com tempo definido:** Pergunta que especifica um período de tempo a ser estudado

**Pergunta enviesada:** Pergunta que direciona as pessoas para uma determinada resposta.

**Pergunta específica:** Pergunta simples, significativa e focada em um único tópico ou em algumas ideias relacionadas entre si

**Pergunta mensurável:** Pergunta cujas respostas podem ser quantificadas e avaliadas

**Pergunta orientada para a ação:** Pergunta cujas respostas levam a uma alteração

**Pergunta parcial:** Pergunta que faz suposições ou que é difícil de se responder de forma honesta

**Pergunta relevante:** Pergunta significativa para o problema a ser solucionado.

**Pixel:** Nas imagens digitais, uma pequena área de iluminação em uma tela de exibição que, ao ser combinada com outras áreas adjacentes, forma uma imagem digital

**Planilha:** Uma planilha digital

**Poder estatístico:** Probabilidade de um teste de significância reconhecer um efeito que está presente

**População:** Em Data Analytics, significa todos os valores de dados possíveis em um conjunto de dados

**Precisão:** Grau em que os dados estão de acordo com a entidade real que está sendo medida ou descrita

**Preenchimento automático:** Boxe no canto inferior à direita na célula selecionada da planilha que pode ser arrastado para células próximas a fim de continuar uma instrução

**Privacidade dos dados:** Preservar as informações de um titular de dados sempre que ocorrer uma transação com os dados

**Processo de análise de dados:** As seis etapas de pergunta, preparação, processamento, análise, compartilhamento e ação, cujo propósito é obter insights que orientem a tomada de decisão informada

**Processo de validação de dados:** Processo de verificação e reavaliação da qualidade dos dados para que sejam completos, precisos, seguros e consistentes

**Proporção:** O princípio do design de usar o tamanho relativo e o arranjo de elementos visuais para expor informações em uma visualização de dados

**Propriedade:** Aspecto da ética de dados que presume que os indivíduos possuem os dados brutos que fornecem e têm controle primário sobre o seu uso, processamento e compartilhamento.

R

**R:** Uma linguagem de programação usada para análise estatística, visualização, e outras análises de dados

**Ranking:** Um sistema para posicionar valores de um banco de dados dentro de uma escala de conquistas ou status

**Realçar tabela:** Uma visualização de dados que usa cor e formatação condicional em uma tabela

**Receita:** Valor total de renda gerada pela venda de produtos ou serviços

**Redundância:** Quando a mesma parte dos dados é armazenada em dois ou mais locais

**Referência absoluta:** Referência em uma função que é bloqueada, para que as linhas e colunas não sofram alterações se a função for copiada

**Referência de célula:** Célula ou intervalo de células em uma planilha, que normalmente se usa em fórmulas e funções

**Reformulação:** Processo de reformular um problema ou desafio e redirecioná-lo para uma resolução possível

**Registro:** Coleção de dados relacionados em uma tabela de dados, geralmente sinônimo de linha

**Regulamento Geral de Proteção de Dados (GDPR, na sigla em inglês):** Entidade que atua na criação de políticas na União Europeia, criada para ajudar a proteger as pessoas e seus dados

**Relatividade:** O processo de considerar as observações em relação - ou em proporção - a algo mais

**Relatório:** Coleção estática de dados entregue periodicamente às partes interessadas.

**Remover duplicações:** Ferramenta de planilha que procura e elimina automaticamente as entradas duplicadas de uma planilha.

**Repetição:** O princípio do design de repetir elementos visuais para expressar significado em uma visualização de dados

**Replicação de dados:** O processo de armazenamento de dados em múltiplos locais

**Repositório de metadados:** Banco de dados criado para armazenar metadados

**Restrições de dados:** Critérios que determinam se um dado é limpo e válido

**Retorno sobre o investimento (ROI):** Fórmula que utiliza métricas de investimento e lucro para avaliar o sucesso de um investimento

**Rhythm:** O princípio do design de criar movimento e fluxo em uma visualização de dados para envolver um público-alvo

**RIGHT JOIN:** Função de SQL que retornará todos os registros da tabela à direita e somente os registros correspondentes da tabela à esquerda

**RIGHT:** Função que retorna um determinado número de caracteres do lado direito de uma string de texto

**ROUND:** Função de SQL que retorna um número arredondado a determinado número de casas decimais

S

**Segurança de dados:** Proteger os dados contra acesso não autorizado ou corrupção por meio da adoção de medidas de segurança

**SELECT INTO:** Uma cláusula SQL que copia dados de uma tabela para uma tabela temporária, sem adicionar a nova tabela no banco de dados

**SELECT:** Seção de uma consulta que indica de qual(is) coluna(s) extrair os dados

**Significância estatística:** Probabilidade de que os resultados da amostra não se devam a um acaso

**Sintaxe:** Estrutura predeterminada de um idioma que inclui todas as palavras, símbolos e pontuação necessários, bem como sua correta colocação

**Sort Range:** Uma função em um menu da planilha que classifica um intervalo especificado e preserva as células fora do intervalo

**Sort Sheet:** Uma função em um menu da planilha que ordena todos os dados pela classificação de uma coluna classificada em específico e mantêm os dados juntos ao longo das linhas

**SPLIT:** Função de planilha que divide o texto em torno de um caractere específico e coloca cada fragmento em uma célula nova e separada

**Spotlightling:** Um recurso que faz uma varredura pelos dados para identificar rapidamente os insights mais importantes.

**SQL:** (Consulte Linguagem de consulta estruturada)

**Storytelling com dados:** Comunicar o significado de uma base de dados com visuais e uma narrativa que são personalizados para um público-alvo em particular

**String de texto:** Grupo de caracteres dentro de uma célula, na maioria das vezes composto de letras

**Subconsulta:** Consulta SQL aninhada em uma consulta maior

**Sublinhados:** Linhas usadas para sublinhar palavras e conectar caracteres de texto

**SUBSTR:** Função SQL que extrai uma substring de uma string variável

**Substring:** Um subconjunto de uma string de texto

**Subtítulo:** Texto que apoia um título ao adicionar contexto e descrição

**SUM:** Função de planilha que soma os valores de um intervalo de células selecionado

**SUMIF:** Função de planilha que adiciona dados numéricos com base em uma condição

**SUMPRODUCT:** Função que multiplica matrizes e retorna a soma desses produtos

T

**Tabela dinâmica:** Ferramenta de resumo de dados utilizada para selecionar, reorganizar, agrupar, contar, calcular o total ou calcular a média dos dados

**Tabela temporária:** Tabela de banco de dados que é criada e existe temporariamente em um servidor de banco de dados

**Tableau:** Uma plataforma de inteligência de negócios e análises que ajuda as pessoas a visualizar, entender, e tomar decisões com dados

**Tarefa de negócios:** Pergunta ou problema que a análise de dados resolve para uma empresa

**Tarefa de resumo:** Tabela usada para resumir informações estatísticas sobre os dados

**Taxa de resposta estimada:** O número médio de pessoas que normalmente completam uma pesquisa

**Taxa de rotatividade:** Taxa que mensura a saída e entrada dos funcionários da empresa

**Teste A/B:** Processo de testar duas variações da mesma página web para determinar qual página é mais bem sucedida em atrair tráfego de usuários e gerar receita

**Teste de hipótese:** Processo para determinar se uma pesquisa ou experimento tem resultados significativos

**Texto alternativo:** Texto que fornece uma alternativa ao conteúdo não textual, tais como imagens e vídeos

**Tipo de dado:** Atributo que descreve uma parte dos dados com base em seus valores, sua linguagem de programação ou as operações que pode realizar

**Tipos de problema:** Os diversos problemas que os analistas de dados encontram, incluindo: categorizar as coisas, descobrir conexões, encontrar padrões, identificar temas, fazer previsões e detectar algo incomum

**Título:** Texto no topo de uma visualização que comunica os dados que estão sendo apresentados

**Tomada de decisão inspirada em dados:** Explorar diferentes fontes de dados para descobrir o que elas têm em comum

**Tomada de decisões com base em dados:** Usar fatos para orientar a estratégia de negócios

**Transferência de dados:** Processo de cópia de dados de um dispositivo de armazenamento para a memória de um computador ou de um computador para outro

**Transparência de transações:** Aspecto da ética de dados que presume que todos os algoritmos e atividades de processamento de dados sejam explicáveis e entendidos pelo indivíduo que fornece os dados

**TRIM:** Função que remove espaços à esquerda, à direita e repetidos nos dados

**Typecasting:** Conversão de dados de um tipo para outro

U

**Único:** Um valor que não pode ter um duplicado

**Unidade:** O princípio do design de usar elementos visuais que se complementam mutuamente para criar apelo estético e clareza em uma visualização de dados

**United States Census Bureau:** Agência do Departamento do Comércio dos EUA que atua como o principal provedor de dados de qualidade do país a respeito de seus habitantes e sua economia

V

**Validação de campo cruzado:** Processo que garante o cumprimento de certas condições para múltiplos campos de dados

**Validação de dados:** Ferramenta para verificar a precisão e a qualidade dos dados

**Validade:** Grau de conformidade dos dados com as restrições quando são inseridos, coletados ou criados

**VALUE:** Função de planilha que converte uma string de texto que representa um número a um valor numérico

**Variedade:** O princípio do design de usar diferentes tipos de elementos visuais em uma visualização de dados para engajar um público-alvo

**Verificação:** Processo para confirmar que uma tentativa de limpeza de dados foi executada com sucesso e que os dados derivados são precisos e confiáveis

**Viés de amostragem:** Representação excessiva ou deficiente de determinados membros de uma população como resultado de trabalhar com uma amostra que não representa a população como um todo

**Viés de confirmação:** Tendência de buscar ou interpretar informações de forma a confirmar crenças existentes

**Viés de dados:** Quando uma preferência a favor ou contra uma pessoa, um grupo de pessoas ou coisas sistematicamente distorce os resultados da análise de dados para uma determinada direção

**Viés de interpretação:** Tendência a interpretar situações ambíguas de forma positiva ou negativa.

**Viés do experimentador:** A tendência de diferentes pessoas observarem as coisas de modo diferente (Consulte viés do observador)

**Viés do observador:** Tendência de diferentes pessoas observarem as coisas de modo diferente (também denominada viés do experimentador)

**Viés:** Uma preferência consciente ou subconsciente que favorece ou desfavorece uma pessoa, um grupo de pessoas ou uma coisa

**Visão circular:** Uma visualização de dados que mostra força comparativa nos dados

**Visualização de dados:** Representação gráfica dos dados

**Visualização estática:** Uma visualização de dados que não muda ao longo do tempo a menos que seja editada

**Visualização:** (Consulte visualização de dados)

**Visualizações dinâmicas:** Visualizações de dados que são interativas ou mudam ao longo do tempo

**VLOOKUP:** Função de planilha que busca verticalmente um determinado valor em uma coluna para retornar uma informação correspondente

W

**WHERE:** Seção de uma consulta que especifica critérios aos quais os dados requisitados devem atender

**WITH:** Uma cláusula SQL que cria uma tabela temporária que pode ser consultada várias vezes