A

**Agenda**: Una vista de las citas programadas.

**Alcance del trabajo (SOW):** Un esquema acordado de las tareas a realizar durante un proyecto.

**Algoritmo:** Proceso o conjunto de reglas a seguir para una tarea específica.

**Análisis computacional de datos:** La ciencia de los datos.

**Análisis de datos:** La recopilación, transformación y organización de los datos para sacar conclusiones, hacer predicciones e impulsar una toma de decisiones fundamentada.

**Análisis de déficits:** Método para examinar y evaluar el estado actual de un proceso con el fin de identificar las oportunidades de mejora en el futuro.

**Analista de datos:** Alguien que recopila, transforma y organiza los datos para sacar conclusiones, hacer predicciones e impulsar la toma de decisiones fundamentada.

**Anonimización de datos:** El proceso de proteger los datos privados o confidenciales de las personas eliminando información que pueda asociarse a ellas.

**Apertura:** El aspecto de la ética de datos que promueve el acceso libre a los datos, su uso y también el uso compartido.

**Archivo CSV (archivo de valores separados por coma):** Archivo de texto delimitado que utiliza una coma para separar los valores.

**Archivo de audio:** Almacenamiento en audio digitalizado generalmente en MP3, AAC u otro formato comprimido.

**Archivo de vídeo:** Conjunto de imágenes, archivos de audio y otros datos generalmente codificados en un formato comprimido, por ejemplo, MP4, MV4, MOV, AVI o FLV.

**Atributo**: Característica o calidad de los datos que se usa para etiquetar una columna en una tabla.

B

**Base de datos normalizada:** Base de datos en la que solo se almacenan datos relacionados en cada tabla.

**Base de datos relacional:** Base de datos que contiene una serie de tablas que se pueden conectar para formar relaciones.

**Base de datos:** Recopilación de datos almacenados en un sistema informático.

**Bloc de notas:** Entorno de programación interactivo y editable para generar informes de datos y mostrar destrezas en el uso de datos.

**Bordes**: Líneas que se pueden agregar alrededor de dos o más celdas en una hoja de cálculo.

C

**Campo:** Información de una fila o columna de una hoja de cálculo; en una tabla de datos, suele ser una columna de la tabla.

**Causa raíz:** La razón por la que ocurre un problema.

**Ciclo de vida de los datos:** Secuencia de etapas por las que pasan los datos, que incluye planificar, capturar, gestionar, analizar, archivar y destruir.

**Ciencia de datos:** Campo de estudio que utiliza datos sin procesar para crear nuevas formas de modelar y entender lo desconocido.

**Clave externa:** Campo en una tabla de una base de datos que constituye una clave primaria en otra tabla (Ver “Clave primaria”).

**Clave primaria:** Identificador en una base de datos que hace referencia a una columna en la que cada valor es único (Ver “Clave externa”).

**Conjunto de datos:** Una colección de datos que pueden ser manipulados o analizados como una unidad.

**Consentimiento:** Aspecto de la ética de datos que presupone el derecho de una persona a conocer cómo y por qué se utilizarán sus datos personales antes de aceptar proporcionarlos.

**Consulta:** Solicitud de datos o información de una base de datos.

**CONTAR**: Función de la hoja de cálculo que cuenta el número de celdas en un rango que cumplen con un criterio específico.

**Contexto:** Condición en la que algo existe o sucede.

**Control de acceso:** Funciones como la protección de contraseñas, permisos de usuario y cifrado que se usan para proteger una hoja de cálculo.

**Controlador de relleno:** Cuadro en la esquina inferior derecha de una celda seleccionada de una hoja de cálculo que se puede arrastrar pasando por celdas vecinas para seguir una instrucción.

**Convenciones de nomenclatura:** Pautas uniformes para el nombre de un archivo que describen el contenido, la fecha de creación y la versión.

**Cookie:** Pequeño archivo almacenado en una computadora que contiene información acerca de sus usuarios.

D

**Datos abiertos:** Datos que están disponibles para el público.

**Datos booleanos:** Tipo de datos con solo dos valores posibles, generalmente verdadero o falso.

**Datos continuos:** Datos que se miden y que pueden tener casi cualquier valor numérico.

**Datos cualitativos:** Medida subjetiva y explicativa de una cualidad o característica.

**Datos cuantitativos:** Medida específica y objetiva, como un número, cantidad o rango.

**Datos de primera fuente:** Datos recopilados por una persona o por un grupo por medio de sus propios recursos.

**Datos de segunda fuente:** Datos que recopila un grupo directamente de su audiencia y que, luego, se venden.

**Datos de terceros:** Datos proporcionados de fuentes externas que no los recopilaron directamente.

**Datos discretos:** Datos que se cuentan y tienen un número limitado de valores.

**Datos en formato ancho:** Conjunto de datos en el que cada tema tiene una sola fila con varias columnas para describir los valores de los distintos atributos del tema.

**Datos en formato largo:** Conjunto de datos en el que cada fila constituye un punto en el tiempo por tema, es decir que cada tema tiene datos en varias filas.

**Datos estructurados:** Datos organizados de cierta forma, por ejemplo, en filas y columnas.

**Datos externos:** Los datos que se alojan y generan fuera de una organización.

**Datos internos:** Datos alojados en los sistemas propios de una empresa.

**Datos no estructurados:** Datos que no se organizan de forma que sea fácil identificarlos.

**Datos nominales:** Tipo de datos cualitativos que se categorizan sin un orden establecido.

**Datos ordinales:** Datos cualitativos con un orden o escala establecidos.

**Datos:** Una colección de hechos.

**DESDE:** Sección de una consulta que indica el origen de los datos seleccionados.

**Destrezas analíticas:** Cualidades y características asociadas al uso de hechos para resolver problemas.

**Diseño de datos:** Cómo se organiza la información.

**Dominio del problema:** Área de análisis que abarca cada actividad que afecta a un problema o se ve afectada por él.

**DÓNDE:** La sección de una consulta que especifica los criterios que deben cumplir los datos solicitados.

E

**Ecosistema de datos:** Los distintos elementos que interactúan entre sí para producir, gestionar, almacenar, organizar, analizar y compartir datos.

**Ecuación:** Cálculo que implica suma, resta, multiplicación o división (también se denomina expresión matemática).

**Elemento de datos:** Cierta información dentro de un conjunto de datos.

**Encabezado:** La primera fila en una hoja de cálculo que hace referencia al tipo de datos en cada columna.

**Equidad:** Cualidad del análisis de datos que no genera sesgos ni los reafirma.

**Esquema:** Forma de describir cómo se organiza algo, por ejemplo, datos.

**Estrategia de datos:** La gestión de las personas, los procesos y las herramientas que se usan en el análisis de datos.

**Ética de datos:** Normas justificadas respecto de lo que está bien y lo que está mal a la hora de recopilar, compartir y usar datos.

**Ética:** Normas justificadas respecto de lo que está bien y lo que está mal. Por lo general, presuponen lo que deben hacer los seres humanos, usualmente en términos de derechos, obligaciones, beneficios para la sociedad, equidad o virtudes específicas.

**Expresión matemática:** Cálculo que implica suma, resta, multiplicación o división (también denominada “ecuación”).

F

**Filtrado:** Proceso que muestra solo los datos que cumplen con un criterio específico mientras oculta el resto.

**Fórmula:** Conjunto de instrucciones que se utilizan para realizar un cálculo utilizando los datos de una hoja de cálculo.

**Foto digital:** Una imagen electrónica o computarizada, generalmente en formato BMP o JPG.

**Fuente de datos correctos:** Fuente de datos confiable, original, integral, actual y citada (ROCCC).

**Fuente de datos erróneos:** Fuente de datos que no es confiable, original, integral, actual ni citada (ROCCC).

**Función matemática**: Función que se utiliza como parte de una fórmula matemática.

**Función:** Comando preestablecido que realiza automáticamente un proceso o tarea especificado utilizando los datos de una hoja de cálculo.

G

**Geolocalización:** La ubicación geográfica de una persona o dispositivo mediante información digital.

**Gobierno de datos:** Proceso para garantizar la gestión formal de los recursos de datos de una empresa.

**Gráfico dinámico:** Gráfico creado a partir de los campos de una tabla dinámica.

H

**Hoja de cálculo:** Una hoja de cálculo digital.

I

**Informe:** Recopilación estática de datos que se entrega periódicamente a los interesados.

**Ingresos:** Cantidad total de ingresos generados por la venta de mercaderías o servicios.

**Interesados:** Personas que invierten tiempo y recursos en un proyecto y se interesan por su resultado.

**Interoperabilidad de los datos:** La capacidad de integrar datos de varias fuentes y un factor clave en el uso satisfactorio de los datos abiertos entre las empresas y los gobiernos.

J

K

L

**Lenguaje de consulta estructurado:** Lenguaje de programación informática utilizado para comunicarse con una base de datos.

**Lenguaje de consulta:** Lenguaje de programación informática utilizado para comunicarse con una base de datos.

M

**Macrodatos:** Conjuntos de datos grandes y complejos que, generalmente, se recopilan durante largos períodos y que permiten que los analistas de datos aborden los problemas comerciales de gran alcance.

**MAX:** Función de la hoja de cálculo que muestra el valor numérico mayor de un rango de celdas.

**Mentalidad técnica:** La capacidad de dividir las cosas en pasos o piezas más pequeñas y trabajar con ellas de forma ordenada y lógica.

**Metadatos administrativos:** Metadato que indica el origen técnico de un recurso digital.

**Metadatos descriptivos:** Metadatos que describen datos y que se pueden utilizar para identificarlos más adelante.

**Metadatos estructurales:** Metadatos que indican cómo se organizan ciertos datos y si forman parte de una recopilación de datos o de varias.

**Metadatos:** Datos sobre datos; en la gestión de bases de datos, ayudan a los analistas de datos a interpretar el contenido de los datos dentro de una base de datos.

**Metodología SMART:** Herramienta para determinar la eficacia de una pregunta según sea específica, mensurable, orientada a la acción, pertinente y con plazos determinados.

**Métrica:** Tipo de datos únicos y cuantificables que se utiliza para medición.

**Microdatos:** Puntos de datos pequeños, específicos, que generalmente involucran un breve período y que son útiles para tomar decisiones diarias.

**MIN**: Función de la hoja de cálculo que muestra el valor numérico menor de un rango de celdas.

**Modelo de datos:** Herramienta para organizar los elementos de datos y la forma en que se relacionan entre ellos.

**Muestra:** En el análisis computacional de datos, segmento de una población que la representa toda.

**Muestreo imparcial:** Muestra de la población que la representa en su totalidad.

N

**Nube:** Lugar para mantener los datos en línea, en lugar de tenerlos en el disco duro de una computadora.

O

**Objetivo medible:** Objetivo medible establecido por una empresa y evaluado mediante métricas.

**Observación:** Los atributos que describen una pieza de datos contenida en una fila de una tabla.

**Oficina del Censo de los Estados Unidos:** Agencia del Departamento de Comercio de los Estados Unidos que funciona como proveedor principal de datos de calidad sobre las personas y la economía a nivel nacional.

**Operador:** Símbolo que designa la operación o cálculo a realizarse.

**Orden de operaciones:** Uso de paréntesis para agrupar los valores de la hoja de cálculo a fin de aclarar el orden en el que deben realizarse las operaciones.

**Ordenación:** El proceso de organizar los datos en un sistema de clasificación significativo para que sean más fáciles de entender, analizar y visualizar

**Organización Mundial de la Salud:** Organización cuya función principal es la de controlar y coordinar la salud a nivel internacional dentro del sistema de las Naciones Unidas.

P

**Panel:** Herramienta que monitorea los datos entrantes en vivo.

**Pensamiento analítico:** El proceso de identificar y definir un problema, para luego resolverlo mediante el uso de datos de manera organizada, paso a paso.

**Pensamiento estructurado:** El proceso de reconocer el problema o la situación actuales, organizar la información disponible, revelar déficits y oportunidades, e identificar opciones.

**Píxel:** En imágenes digitales, es un área pequeña de iluminación en una pantalla de visualización que, cuando se combina con otras áreas adyacentes, forma una imagen digital.

**Población:** En el análisis computacional de datos, todos los valores posibles en un conjunto de datos.

**Pregunta con límite de tiempo:** Pregunta que especifica un plazo a analizar.

**Pregunta específica:** Pregunta simple, significativa y enfocada en un solo tema o en algunas ideas estrechamente relacionadas entre sí.

**Pregunta injusta:** Pregunta en la que se hacen suposiciones o que es difícil de responder honestamente.

**Pregunta medible:** Pregunta cuyas respuestas se pueden cuantificar y evaluar.

**Pregunta orientada a la acción:** Pregunta cuyas respuestas conducen al cambio.

**Pregunta principal:** Pregunta que orienta a las personas hacia cierta respuesta.

**Pregunta relevante:** Pregunta que tiene importancia para el problema que se debe resolver.

**Privacidad de los datos:** Preservación de la información sobre los datos de una persona cada vez que ocurre una transacción de datos.

**Proceso de análisis de datos:** Las seis fases de preguntar, preparar, procesar, analizar, compartir y actuar cuyo propósito es el de obtener conocimiento que propicie la toma de decisiones informada.

**PROMEDIO**: Función de una hoja de cálculo que muestra el resultado de un promedio de los valores de un rango seleccionado.

**Propiedad**: El aspecto de la ética de datos que presupone que cada persona es dueña de los datos sin procesar que proporciona y que tiene control primordial sobre su uso, procesamiento y uso compartido.

Q

R

**Rango:** Conjunto de dos o más celdas en una hoja de cálculo.

**Redes sociales:** Sitios web y aplicaciones en los que los usuarios crean y comparten contenido o interactúan entre sí.

**Redundancia:** Cuando los mismos datos se almacenan en dos o más lugares.

**Referencia de celda:** Una celda o un rango de celdas en una hoja de cálculo que se utiliza generalmente en las fórmulas y funciones.

**Reformulación:** Proceso de replantear un problema o desafío, que se redirecciona luego hacia una posible resolución.

**Registro:** Conjunto de datos relacionados en una tabla de datos, generalmente sinónimo de fila.

**Reglamento general de protección de datos de la Unión Europea (GDPR):** Organismo formulador de políticas en la Unión Europea, creado para ayudar a proteger a las personas y sus datos.

**Repositorio de metadatos:** Base de datos creada para almacenar metadatos.

**Retorno de la inversión (ROI):** Fórmula que utiliza las métricas de inversión y ganancias para evaluar el éxito de una inversión.

S

**Seguridad de datos:** Emplear medidas de seguridad para proteger los datos contra el acceso no autorizado o contra la corrupción.

**SELECCIONAR:** Sección de una consulta que indica el subconjunto de un conjunto de datos.

**Sesgo de confirmación:** Tendencia a buscar o interpretar la información de manera que confirme creencias preexistentes.

**Sesgo de interpretación:** Tendencia a interpretar situaciones ambiguas de manera positiva o negativa.

**Sesgo de los datos:** Cuando una preferencia a favor o en contra de una persona, un grupo de personas o una cosa sesga sistemáticamente los resultados del análisis de datos en una cierta dirección.

**Sesgo del investigador:** Tendencia de distintas personas a observar las cosas de forma diferente (Ver “Sesgo del observador”).

**Sesgo del muestreo:** Representar en mayor o en menor medida a ciertos miembros de una población debido a que se trabaja con una muestra que no representa a la población en su totalidad.

**Sesgo del observador:** Tendencia de distintas personas a observar las cosas de forma diferente (también se denomina “sesgo del investigador”).

**Sesgo:** Preferencia consciente o no a favor o en contra de una persona, un grupo de personas o una cosa.

**SQL:** (Ver “Lenguaje de consulta estructurado”).

**SUMA:** Función de una hoja de cálculo que suma los valores de un rango de celdas seleccionadas.

T

**Tabla dinámica:** Herramienta de resumen de datos que se utiliza para clasificar, reorganizar, agrupar, contar, totalizar o promediar datos.

**Tarea empresarial:** La pregunta o el problema que el análisis de datos resuelve para un negocio.

**Tasa de rotación:** Ritmo en el que los empleados abandonan voluntariamente una empresa.

**Tipo de datos de texto:** Secuencia de caracteres y puntuación que contiene información textual (también denominado “tipo de datos de cadena”).

**Tipo de datos en cadena:** Secuencia de caracteres y puntuación que contiene información textual (también denominado “tipo de datos de texto”).

**Tipo de datos:** Un atributo que describe cierto dato según sus valores, su lenguaje de programación o las operaciones que puede realizar.

**Tipos de problemas:** Distintos problemas a los que se enfrenta el analista de datos, que incluyen categorizar elementos, descubrir conexiones, hallar patrones, identificar temas, hacer predicciones y detectar algo inusual.

**Toma de decisiones basada en datos:** Uso de datos para guiar la estrategia empresarial.

**Toma de decisiones inspirada en datos:** Exploración de diferentes fuentes de datos para descubrir qué tienen en común.

**Transparencia de la transacción:** Aspecto de la ética de datos que presupone que se deben explicar todas las actividades de procesamiento de datos y los algoritmos a la persona que proporciona los datos y que también presupone que esta persona debe comprenderlos.

U

V

**Vigencia:** El aspecto de la ética de datos que presupone que las personas deben conocer las transacciones financieras resultantes del uso de sus datos personales y la magnitud de esas transacciones.

**Visualización de datos:** La representación gráfica de los datos.

**Visualización:** (Ver “Visualización de datos”).

W

X

Y

Z