

CIENCIA DE DATOS

Área personal / Mis cursos / [CIENCIA DE DATOS](#) / [UNIDAD 2: ANALÍTICA DE DATOS](#)
/ [FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y S...](#)



Buscar en los foros



FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

Vencimiento: Monday, 29 de May de 2023, 23:50

✓ Hecho

⚙️ Configuraciones ▾

Mostrar respuestas anidadas

Se ha alcanzado la fecha límite para publicar en este foro, por lo que ya no puede publicar en él.

FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

Monday, 29 de May de 2023, 18:30

[Enlace permanente](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
de [HENRY ANTONIO JARAMILLO OLIVARES](#) - Monday, 29 de May de 2023, 20:12

Los [conceptos de distancia y similitud](#) son fundamentales para analizar y comparar conjuntos de datos. La distancia se utiliza para medir la separación o diferencia entre dos muestras o puntos de datos, mientras que la similitud busca establecer qué tan parecidos son.

La distancia euclídea es ampliamente utilizada en ciencia de datos para medir la distancia entre puntos en un espacio multidimensional.

Por ejemplo, supongamos que tienes una lista de reproducción de música A con las canciones ["Canción 1", "Canción 2", "Canción 3"] y otra lista B con las canciones ["Canción 2", "Canción 4", "Canción 5"]. Puedes calcular la distancia entre estas dos listas contando cuántas canciones son diferentes en total. En este caso, la distancia sería 4, ya que hay cuatro canciones que son diferentes entre ambas listas.

Por otro lado, la similitud se refiere a la medida de qué tan parecidos son dos conjuntos de datos. Por ejemplo, imagina que tienes una base de datos de usuarios y quieres encontrar usuarios similares en función de sus intereses. Tienes dos usuarios, A y B, y sus intereses son los siguientes:

Usuario A: ["Deportes", "Música", "Viajes"]

Usuario B: ["Música", "Cine", "Libros"]

Para medir la similitud entre estos usuarios, puedes utilizar el enfoque de coincidencia de intereses. En este caso, la similitud sería 1, ya que ambos usuarios tienen "Música" en común, a pesar de que tienen otros intereses diferentes. Esto indica que son similares en términos de su interés compartido en la música.

Estos son conceptos clave para medir y comparar conjuntos de datos. La distancia cuantifica la diferencia entre puntos o muestras de datos, mientras que la similitud busca establecer qué tan parecidos son. Estas medidas son fundamentales para diversas tareas de análisis y modelado de datos en la ciencia de datos.

Como Tema de Muestra de Datos se tiene: Registro de ventas de una tienda de productos el cual registra:

- ID Cliente (número identificador del cliente)
- Zona
- País
- Tipo de producto
- Canal de venta
- Prioridad

- Fecha de pedido
- ID Pedido
- Fecha de envío
- Unidades
- Precio unitario
- Coste unitario
- Importe de venta
- Importe de coste

[Enlace permanente](#)[Mostrar mensaje anterior](#)**Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD**de [ANDRICK STEVEN VIZUETA LOPEZ](#) - Monday, 29 de May de 2023, 20:15

Los [conceptos de distancia y similitud](#) son importantes en la ciencia de datos y pueden ser útiles en Microsoft Power BI. La distancia se refiere a la medida de la diferencia entre dos objetos o puntos de datos, mientras que la similitud mide qué tan parecidos son dos objetos o puntos de datos. Un ejemplo práctico en Microsoft Power BI sería utilizar la función de clustering para analizar la similitud entre diferentes puntos de datos en un conjunto de datos. Por ejemplo, si estamos trabajando con un conjunto de datos de ventas de diferentes productos, podríamos utilizar la función de clustering para agrupar los productos en función de su similitud en términos de características de ventas, como la cantidad vendida, el margen de beneficio y el precio de venta.

[Enlace permanente](#)[Mostrar mensaje anterior](#)**Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD**de [ANDRICK STEVEN VIZUETA LOPEZ](#) - Monday, 29 de May de 2023, 21:34

https://drive.google.com/file/d/1DvGyj2d6s4lnCGB2wR0pj09rNloBYoJq/view?usp=drive_link

[Enlace permanente](#)[Mostrar mensaje anterior](#)**Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD**de [GABRIELA SARAY QUIMIS ESPINOZA](#) - Monday, 29 de May de 2023, 20:16

Los científicos de datos usan análisis de distancia y similitud para visualizar patrones y relaciones entre variables. La similitud mide su proximidad o semejanza, la distancia mide la diferencia entre dos objetos o puntos. Para explorar y visualizar patrones y relaciones entre objetos, las técnicas estadísticas y de análisis de datos utilizan ambos conceptos.

- Las mediciones de **distancia** son fáciles de comprender y calcular, pero es posible que no capturen con precisión las similitudes entre las observaciones realizadas utilizando varias escalas o unidades.
- Las medidas de **similitud** son útiles para determinar qué tan estrechamente relacionadas están las observaciones y pueden revelar información sobre la dirección de las relaciones, pero pueden ser más difíciles de calcular y comprender.

En ingeniería de sistemas se aplica estos conceptos para la agrupación de datos, recomendaciones de productos, análisis de redes sociales, reconocimiento de patrones y análisis de sentimientos, entre otras áreas.

Mi tema a analizar será: Tasa de natalidad, nacidos vivos en un año (por cada 1.000 personas) mundialmente

Enlace de base de datos: https://datos.bancomundial.org/indicator/SP.DYN.CBRT.IN?end=2021&name_desc=false&start=1960&view=chart

[Enlace permanente](#)[Mostrar mensaje anterior](#)**Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD**de [ANGELO DAVID PAZMIÑO QUIMI](#) - Monday, 29 de May de 2023, 20:19

La distancia y la similitud son ideas esenciales utilizadas para medir la proximidad o la relación entre varios objetos o conjuntos de datos en el campo de la ciencia de datos.

Una métrica numérica llamada distancia muestra la distancia o diferencia entre dos elementos. En la ciencia de datos se emplean diferentes medidas de distancia en función del tipo de datos y del contexto del problema.

Mientras que la similitud es una métrica que muestra lo parecidos o cercanos que están dos objetos o conjuntos de datos. Los científicos de datos pueden elegir entre diversas métricas de similitud en función del contexto del problema y del tipo de datos que se utilicen.

Mi tema a analizar es: Conjunto de datos del censo de población del estado de Maharashtra, India.

Base de datos: <https://www.kaggle.com/datasets/prasad22/mha-population-dataset>

[Enlace permanente](#)[Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
de [KAYSY MARCELA BARRERA PEREZ](#) - Monday, 29 de May de 2023, 20:20

La distancia es una medida cuantitativa que indica la separación o diferencia entre dos elementos o puntos en un espacio. Algunas medidas de distancia comunes son la distancia euclídea, la distancia Manhattan y la distancia de Minkowski.

La similitud es una medida que indica cuán cercanos o parecidos son dos elementos o conjuntos de datos. Algunas medidas de similitud comunes son la similitud coseno, el coeficiente de correlación de Pearson y la distancia de Jaccard.

[Enlace permanente](#) [Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
de [JOSE CARLOS PAYE ARROYO](#) - Monday, 29 de May de 2023, 20:22

En ciencia de datos, tanto la distancia como la similitud son conceptos fundamentales para comparar y medir la relación entre objetos o muestras. Estos conceptos se utilizan en una variedad de áreas, como la clasificación, la agrupación (clustering), la recomendación y el reconocimiento de patrones, entre otros.

La distancia es una medida numérica que cuantifica la diferencia entre dos objetos. Se utiliza para determinar qué tan lejos o cercanos están dos puntos en un espacio de características o atributos. Existen diferentes métricas de distancia que se pueden utilizar dependiendo del contexto y de las características de los datos.

La similitud, por otro lado, es una medida que cuantifica qué tan parecidos son dos objetos o muestras. A diferencia de la distancia, la similitud se define de tal manera que valores más altos indican una mayor semejanza entre los objetos. Por lo tanto, la similitud es inversamente proporcional a la distancia. Las medidas de similitud son ampliamente utilizadas en técnicas de comparación, búsqueda de información y filtrado colaborativo.

Tema: Esperanza de vida en Ecuador comparación por género.
-> <https://datosmacro.expansion.com/demografia/esperanza-vida/ecuador>

[Enlace permanente](#) [Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
de [FERNANDO ANDRE SANCHEZ MOREIRA](#) - Monday, 29 de May de 2023, 20:24

Distancia: La distancia es una medida que cuantifica la separación o la diferencia entre dos elementos o puntos en un espacio. En el contexto de la ciencia de datos, la distancia se utiliza para determinar la similitud o la diferencia entre objetos o conjuntos de datos. Similitud: La similitud es una medida que cuantifica la semejanza o la relación entre dos elementos o conjuntos de datos. Mientras que la distancia mide la separación o la diferencia, la similitud mide el grado de correspondencia o proximidad entre ellos.

El tema para analizar la siguiente clase será:
Student Mental health
Enlace: <https://www.kaggle.com/datasets/shariful07/student-mental-health>

[Enlace permanente](#) [Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
de [NICOLE ANGELA HOLGUIN SANCAN](#) - Monday, 29 de May de 2023, 20:28

En la ciencia de datos los [conceptos de distancia y similitud](#) se determinan como importantes en el análisis y la clasificación de datos.

En primer lugar, la distancia se refiere a una medida numérica que indica qué tan diferentes o cercanos están dos puntos o conjuntos de datos. Permite identificar patrones, agrupar elementos similares y calcular la similitud entre ellos.

Por otro lado, la similitud es una medida que determina qué tan parecidos son dos puntos o conjuntos de datos. Tanto la distancia como la similitud son herramientas esenciales para la detección de anomalías, el aprendizaje automático y la toma de decisiones basadas en datos.

Mi tema a analizar es el salario de un grupo de personas con distintos títulos profesionales, nivel educativo y años de experiencia.

Enlace de base de datos: <https://www.kaggle.com/datasets/mohithsairamreddy/salary-data>

[Enlace permanente](#) [Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
de [ANTHONY ELIAN MONCAYO FAJARDO](#) - Monday, 29 de May de 2023, 20:31

Dentro de la ciencia de datos la cual se centra en el análisis y la interpretación de grandes volúmenes de datos para poder obtener conocimiento significativo. Los [conceptos de distancia y similitud](#) son fundamentales para la comprensión y el desarrollo de diversos algoritmos con técnicas que se utilizan en dichos análisis. La distancia es una medida cuantitativa en la cual es la separación entre dos objetos o puntos en un espacio multidimensional.

Los conceptos de similitud y distancia son fundamentales en la ciencia de datos para poder evaluar la relación y la semejanza entre puntos de datos. Las medidas de distancia, se utilizan para calcular la diferencia entre puntos a comparación de la similitud, en la cual se usan para evaluar la similitud en una agrupación de datos o clustering.

Tema a analizar es: Redes sociales con mayor numero de usuarios activos a nivel mundo

Enlace: <https://es.statista.com/estadisticas/600712/ranking-mundial-de-redes-sociales-por-numero-de-usuarios/>

[Enlace permanente](#) [Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
de [PEDRO ANDRES DELGADO FLORES](#) - Monday, 29 de May de 2023, 20:32

La distancia y la similitud son variables cuantitativa que sirven para el análisis estadístico de los datos, para la distancia mide la diferencia entre dos objetos de datos y esto permite emplear algoritmos y aplicaciones para el análisis de datos, mientras que la similitud mide el grado de semejanza que tienen dos objetos de datos y esto permite establecer técnicas para entender la relación de los datos. Se los puede aplicar para la agrupación de datos que nos permite identificar patrones en los datos, el análisis de las redes sociales que nos permite reconocer comportamiento entre las interacciones que tienen las personas mientras usan las redes sociales, el reconocimiento de productos para saber las similitudes y características que tiene un producto con otro para establecer productos que le sea de mayor interés para el cliente, la recuperación de documentos que establece la distancia que se toma en buscar un documento entre otro.

Tema: Global Internet Usage (Uso de internet en el mundo)

<https://www.kaggle.com/datasets/sansuthi/gapminder-internet?resource=download>

[Enlace permanente](#) [Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
de [FREDDY GREGORY URETA VARGAS](#) - Monday, 29 de May de 2023, 20:32

La distancia y la similitud son conceptos fundamentales en la ciencia de datos para comparar y analizar objetos. La distancia es una medida cuantitativa que representa la separación o diferencia entre dos objetos o puntos en un espacio. Permite calcular la magnitud de la divergencia entre ellos, utilizando diversas métricas dependiendo del contexto y el tipo de datos involucrados. Por otro lado, la similitud es una medida que evalúa el grado de semejanza o relación entre dos objetos. A diferencia de la distancia, que se enfoca en medir la diferencia, la similitud busca cuantificar la cercanía o parecido entre los objetos. Estos conceptos son ampliamente utilizados en ciencia de datos para tareas como clasificación, agrupamiento, recomendación y búsqueda de información relevante, permitiendo comparar y relacionar datos de manera objetiva y cuantitativa.

Tema para analizar: Datos de la población mundial 1960-2020

Link: <https://www.kaggle.com/datasets/utkarshx27/world-population-data-1960-2020>

[Enlace permanente](#) [Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
de [ISAAC ROBERTO GOMEZ JUNCO](#) - Monday, 29 de May de 2023, 20:36

La distancia se refiere a la separación o diferencia entre dos objetos en un espacio multidimensional. Imagina que tienes un conjunto de puntos dispersos en un plano o en un espacio tridimensional. La distancia nos permite medir la "lejanía" entre dos puntos en ese espacio. Es una medida cuantitativa que nos ayuda a comprender la magnitud de la diferencia entre dos objetos. Las diferentes métricas de distancia, como la euclídea, la de Manhattan o la de Minkowski, nos proporcionan formas de calcular esa distancia en función de las características y dimensiones de los objetos que estamos comparando.

Por otro lado, la similitud se refiere al grado de semejanza o cercanía entre dos objetos. A diferencia de la distancia, la similitud tiende a ser más alta cuanto más cercanos o similares son los objetos. Es una medida que nos permite evaluar cuánto se asemejan dos objetos en términos de características o propiedades específicas. Cuanto mayor sea la similitud, mayor será el grado de

correspondencia o concordancia entre los objetos. Las medidas de similitud, como el coeficiente de similitud de Jaccard o el coeficiente de similitud de coseno, nos ayudan a calcular numéricamente esa similitud en función de las características o atributos que estamos considerando.

Tema a analizar:

Acceso a la electricidad (% de la población)

<https://data.worldbank.org/indicator/EG.ELC.ACCTS.ZS>

[Enlace permanente](#)

[Mostrar mensaje anterior](#)

Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [KARELYS ALEXANDRA BOHORQUEZ DUARTE](#) - Monday, 29 de May de 2023, 20:36

La distancia y la similitud son dos conceptos básicos en el análisis de datos y las estadísticas.

La distancia es una medida cuantitativa de la diferencia entre dos objetos o puntos en una habitación, estas métricas se utilizan para medir la distancia entre dos puntos en diferentes contextos, como en el cálculo de la similitud entre dos textos o la agrupación de objetos en un espacio multidimensional.

La similitud, por otro lado, es una medida cuantitativa de la proximidad o similitud entre dos objetos o puntos, la similitud es una medida inversa de la distancia y se puede definir de diversas maneras, como la similitud coseno o la similitud de Jaccard. La similitud se utiliza en muchas aplicaciones de análisis de datos, como la recuperación de información, la recomendación de productos, la clasificación de imágenes y la agrupación de datos.

La distancia y similitud se aplican en diversos campos de la ingeniería en sistemas. Algunas de las aplicaciones comunes son:

- Agrupación de datos
- Análisis de redes sociales
- Reconocimiento de patrones

Mi base de datos es: Datos de ventas de ejemplo

Enlace: <https://buscarv.com/plantillas/listados-con-datos-de-ejemplo-para-descargar/>

[Enlace permanente](#)

[Mostrar mensaje anterior](#)

Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [VILMA PATRICIA RAMIREZ AGILA](#) - Monday, 29 de May de 2023, 20:38

La similitud se refiere a la medida de cuán parecidos son dos elementos, mientras que la distancia se refiere a la medida de cuán diferentes son. En términos generales, cuanto mayor es la similitud, menor es la distancia.

Mi tema a analizar es el conjunto de datos del modelo de automóvil 2023

Enlace de base de datos: <https://www.kaggle.com/datasets/peshimaammuzammil/2023-car-model-dataset-all-data-you-need>

[Enlace permanente](#)

[Mostrar mensaje anterior](#)

Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [ARTURO ENRIQUE ROSERO ESPINOZA](#) - Monday, 29 de May de 2023, 20:38

La distancia se refiere a la separación o diferencia entre dos objetos en un espacio multidimensional. Imagina que tienes un conjunto de puntos dispersos en un plano o en un espacio tridimensional. La distancia nos permite medir la "lejanía" entre dos puntos en ese espacio. Es una medida cuantitativa que nos ayuda a comprender la magnitud de la diferencia entre dos objetos. Las diferentes métricas de distancia, como la euclídea, la de Manhattan o la de Minkowski, nos proporcionan formas de calcular esa distancia en función de las características y dimensiones de los objetos que estamos comparando.

Por otro lado, la similitud se refiere al grado de semejanza o cercanía entre dos objetos. A diferencia de la distancia, la similitud tiende a ser más alta cuanto más cercanos o similares son los objetos. Es una medida que nos permite evaluar cuánto se asemejan dos objetos en términos de características o propiedades específicas. Cuanto mayor sea la similitud, mayor será el grado de correspondencia o concordancia entre los objetos. Las medidas de similitud, como el coeficiente de similitud de Jaccard o el coeficiente de similitud de coseno, nos ayudan a calcular numéricamente esa similitud en función de las características o atributos que estamos considerando.

[Enlace permanente](#)

[Mostrar mensaje anterior](#)

Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [ALBA DENIS RUIZ GUAMAN](#) - Monday, 29 de May de 2023, 20:39

Los [conceptos de distancia y similitud](#) se basan en su importancia como herramientas que se utilizan comúnmente para visualizar patrones de relaciones entre variables también son muy útiles para explorar datos de redes sociales.

Cuando hemos creado una matriz de similitud o distancia describiendo todas las parejas de actores, podemos estudiar la similitud de las diferencias entre las relaciones de "casos" de la misma manera que estudiariamos las similitudes entre atributos.

[Enlace permanente](#) [Mostrar mensaje anterior](#)

Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [OLIVER MICHAEL TUBAY ZAMBRANO](#) - Monday, 29 de May de 2023, 20:41

La distancia y la similitud son conceptos fundamentales en el campo del análisis de datos. Estas medidas nos permiten comparar y evaluar la relación entre pares de objetos de datos, ya sean especies en biología, usuarios en sistemas de recomendación o documentos en [minería de texto](#). Es importante destacar que la forma en que calculamos la distancia y la similitud depende de la escala de medición de los atributos o características que estamos considerando. Estos atributos pueden ser de naturaleza binaria, cualitativa o cuantitativa. Por ejemplo, si estamos comparando especies de animales, podríamos utilizar atributos binarios para representar la presencia o ausencia de ciertas características. La distancia y la similitud son herramientas versátiles que se utilizan en diversas técnicas de análisis de datos. Por ejemplo, en la clasificación, estas medidas se emplean para agrupar objetos de datos similares en categorías predefinidas.

Tema a analizar: Suicidio global, Salud mental, Uso de sustancias

Link: <https://www.kaggle.com/datasets/thedevastator/global-suicide-mental-health-substance-use-disorder>

[Enlace permanente](#) [Mostrar mensaje anterior](#)

Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [CZISKA WALESKA MORAN ARMIJOS](#) - Monday, 29 de May de 2023, 20:45

Distancia:

La distancia es una medida numérica que representa la diferencia o la separación entre dos elementos o conjuntos de datos. Se utiliza para cuantificar la disimilitud entre ellos. Por lo general, cuanto mayor sea el valor de la distancia, mayor será la diferencia entre los elementos. Una distancia de cero indica que los elementos son idénticos. Las distancias pueden tener diferentes escalas y unidades, dependiendo del contexto y la medida utilizada.

Similitud:

La similitud es una medida que indica qué tan parecidos o relacionados están dos elementos o conjuntos de datos. Se utiliza para cuantificar la semejanza entre ellos. A diferencia de la distancia, en este caso, cuanto mayor sea el valor de similitud, mayor será el grado de parecido entre los elementos. Un valor de similitud de 1 indica que los elementos son idénticos, mientras que un valor de 0 indica que no tienen ninguna similitud. Las medidas de similitud pueden variar según el dominio y el tipo de datos.

DATA:

Mortalidad provocada por lesiones por accidentes de tránsito (por cada 100 000 personas)

<https://datos.bancomundial.org/indicador/SH.STA.TRAF.P5>

[Enlace permanente](#) [Mostrar mensaje anterior](#)

Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [AARON JOEL ACOSTA MONTIEL](#) - Monday, 29 de May de 2023, 20:45

La distancia y la similitud son dos conceptos básicos y a continuación un resumen de ello:

La distancia: Es una medida cuantitativa de la diferencia entre dos objetos o puntos en una habitación, estas métricas se utilizan para medir la distancia entre dos puntos en diferentes contextos, como en el cálculo de la similitud entre dos textos o la agrupación de objetos en un espacio multidimensional.

La similitud : Es una medida cuantitativa de la proximidad o similitud entre dos objetos o puntos, la similitud es una medida inversa de la distancia y se puede definir de diversas maneras, como la similitud coseno o la similitud de Jaccard. La similitud se utiliza en muchas aplicaciones de análisis de datos, como la recuperación de información, la recomendación de productos, la clasificación de imágenes y la agrupación de datos.

Mi tema a analizar es: Videojuegos (ERD y SQL)

Base de datos: <https://www.databasestar.com/sample-database-video-games/>

[Enlace permanente](#)

[Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [CARLOS DAVID GARCIA CEDEÑO](#) - Monday, 29 de May de 2023, 20:49

Distancia y similitud son conceptos fundamentales en diversos campos, como la estadística, la informática, la inteligencia artificial y la ciencia de datos. La distancia es una medida cuantitativa que indica la separación o diferencia entre dos objetos, elementos o puntos en un espacio. Se utiliza para comparar la similitud o diferencia entre ellos. La similitud es una medida que indica qué tan parecidos o cercanos son dos objetos, elementos o puntos en un espacio. A diferencia de la distancia, la similitud generalmente es una medida no negativa que cuantifica la cercanía o relación entre ellos.

Tema para analizar: Venta de licores Iowa clase "E" por producto y fecha de compra para el año 2021.

Link: [2021 Iowa Liquor Sales | data.iowa.gov](https://data.iowa.gov/dataset/2021-Iowa-Liquor-Sales)

[Enlace permanente](#)

[Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [DARLY YENEDY MORAN ESTUPIÑAN](#) - Monday, 29 de May de 2023, 20:49

Distancia:

*En matemáticas y estadísticas, la distancia es una medida cuantitativa que expresa la separación entre dos objetos o puntos en un espacio. Puede haber diferentes tipos de distancias, como la distancia euclídea, la distancia de Manhattan, la distancia de Minkowski, etc.

*La distancia euclídea es la medida más común y se basa en el teorema de Pitágoras. Calcula la longitud del segmento de línea recta que conecta dos puntos en un espacio euclídeo.

*La distancia de Manhattan, también conocida como distancia de la ciudad o distancia L1, mide la suma de las diferencias absolutas entre las coordenadas de dos puntos. Se utiliza frecuentemente en entornos urbanos donde el movimiento solo se permite en direcciones ortogonales.

*La distancia de Minkowski generaliza tanto la distancia euclídea como la distancia de Manhattan y se define como la raíz n-ésima de la suma de las diferencias elevadas a la potencia n entre las coordenadas de dos puntos.

Similitud:

*La similitud es una medida de comparación entre dos objetos o conjuntos de datos. Cuanto mayor sea la similitud, mayor será la semejanza o consecuencias entre los objetos.

*La similitud puede ser evaluada utilizando diferentes técnicas y métodos dependiendo del dominio y los datos específicos. Algunas medidas comunes de similitud incluyen la similitud del coseno, la similitud de Jaccard y la similitud de conexiones.

*La similitud de Jaccard se utiliza para evaluar la similitud entre dos conjuntos. Se calcula dividiendo el tamaño de la intersección de los conjuntos entre el tamaño de la unión de los conjuntos.

Estos [conceptos de distancia y similitud](#) son fundamentales en campos como la estadística, el aprendizaje automático, la minería de datos y la recuperación de información, donde se requiere comparar y evaluar la relación entre objetos o conjuntos de datos.

Tema: Desempleo, mujeres

Link: https://datos.bancomundial.org/indicador/SL.UEM.TOTL.FE.ZS?name_desc=false

[Enlace permanente](#)

[Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [DAHIANA LISSI TIERRA QUINTO](#) - Monday, 29 de May de 2023, 20:52

Dentro de la ciencia de datos la distancia y similitud son fundamentales, puesto que permiten comparar y medir la relación entre objetos o conjunto de datos. Por ello, es importante conocer sus conceptos. Por un lado, la distancia es una medida numérica que cuantifica la separación entre dos objetos o puntos en un espacio de datos, puede ser medida dependiendo de su contexto y tipo de datos. Por otro lado, la similitud es una medida que cuantifica qué tan parecidos son dos objetos o conjuntos de datos. A diferencia de la distancia, donde una mayor distancia indica mayor separación, en la similitud una mayor similitud implica mayor cercanía o parecido. Estos conceptos tienen diversas aplicaciones en la ingeniería por ende juegan un papel importante dentro de la misma y de la ciencia de datos.

Caso de estudio a analizar: Migración neta – Ecuador

<https://datos.bancomundial.org/indicador/SM.POP.NETM?locations=EC>

[Enlace permanente](#)

[Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
de [XAVIER ROBERTO CRUZ LADINES](#) - Monday, 29 de May de 2023, 21:02

En el ámbito de la ciencia de datos, la distancia es una medida que cuantifica la separación o diferencia entre dos conjuntos de datos. Se utiliza para determinar la distancia o disimilitud entre puntos individuales o conjuntos completos de datos. Las métricas de distancia, como la distancia Euclídea o la distancia de Manhattan, proporcionan una forma de calcular esta separación de acuerdo con la naturaleza y estructura de los datos.

La similitud se emplea para evaluar qué tan parecidos o cercanos están dos conjuntos de datos entre sí. En lugar de enfocarse en la diferencia, la similitud resalta las características compartidas o la relación entre los elementos. Se utilizan técnicas como el coeficiente de correlación, la similitud coseno o el índice de Jaccard para calcular la similitud en función de la naturaleza de los datos y los objetivos del análisis.

Tema: "Datos de vuelos y aerolíneas"

Link: <https://www.kaggle.com/usdot/flight-delays>

[Enlace permanente](#)

[Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
de [ADRIANA GEOVANINA VELASTEGUI SANDOVAL](#) - Monday, 29 de May de 2023, 21:07

Los temas distancia y similitud son fundamentales para comprender y medir la relación entre diferentes objetos o puntos de datos. Estos conceptos proporcionan herramientas cuantitativas para evaluar la proximidad o la diferencia entre elementos.

La distancia se refiere a una medida cuantitativa de separación o diferencia entre dos elementos en un espacio multidimensional. Es una forma de cuantificar cómo de lejos o cercanos están dos puntos en función de sus atributos o características.

Por otro lado, la similitud es una medida que evalúa qué tan parecidos o relacionados están dos elementos entre sí. En contraste con la distancia, la similitud busca cuantificar la proximidad o semejanza entre elementos en lugar de su separación. La similitud se utiliza para encontrar elementos que comparten características comunes o para identificar patrones y relaciones.

Mi tema a analizar es:

<https://www.kaggle.com/datasets/adritpal08/indian-stock-market-dataset>

[Enlace permanente](#)

[Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
de [NAYELI SOLIS CHERE](#) - Monday, 29 de May de 2023, 21:07

Las [técnicas de similitud y distancia](#) son fundamentales en la ciencia de datos para comparar y medir la distancia o similitud entre diferentes objetos o instancias.

Las técnicas de distancia miden de forma cuantitativa la diferencia entre dos objetos o puntos. Son fáciles de entender y calcular lo que facilita la realización de gráficos. Entre las medidas de distancia más comunes están: euclidiana, la eliminación de Manhattan, la distancia entre Minkowski y Mahalanobis.

Las técnicas de similitud miden de igual forma, pero en este caso la proximidad o similitud entre dos puntos u objetos. Son empleadas para calcular qué tan cerca están los puntos entre sí, referente a la cercanía o simetría. Entre las más comunes se encuentran: la similitud del coseno, la similitud del Jaccard y la similitud de Pearson. Ambas se expresan mediante números.

Las técnicas de distancia pueden emplearse para medir la similitud debido a que mientras menos diferencia haya entre los objetos, más similitud habrá entre los mismos y viceversa con las técnicas que similitud.

Estas técnicas pueden ser aplicadas de diferentes formas en la ciencia de datos como en la identificación de patrones, recomendación de productos, en la minería de datos, entre otros.

Caso a analizar: Inmunización contra el sarampión (% de niños entre 12 y 23 meses de edad)

Link: <https://datos.bancmundial.org/indicador/SH.IMM.MEAS?locations=1W&start=1983&end=2018&view=chart>

[Enlace permanente](#)

[Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
de [OSCAR JOEL MORAN CEDEÑO](#) - Monday, 29 de May de 2023, 21:09

En la ciencia de datos, la similitud se refiere a qué tan similares son dos puntos de datos entre sí. Es una medida que evalúa la semejanza, cercanía o correspondencia entre las características o atributos de los puntos de datos. La similitud puede ser cuantificada utilizando diferentes métodos y medidas, como la similitud de Jaccard o la similitud de coseno.

La distancia en la ciencia de datos se refiere a la medida de separación o diferencia entre dos puntos de datos. Es una medida que cuantifica la magnitud del espacio entre los puntos y puede ser calculada utilizando diversas métricas, como la distancia euclídea, la distancia Manhattan o la distancia de Minkowski.

Datos a analizar de la: Encuesta Nacional de Empleo, Desempleo y Subempleo - anual 2022, que permite tener indicadores de

empleo a nivel nacional

link del origen de datos: <https://aplicaciones3.ecuadorencifras.gob.ec/BIINEC-war/index.xhtml> (Estadísticas sociodemográficas y sociales- sección trabajo)

[Enlace permanente](#)

[Mostrar mensaje anterior](#)

Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [ALEXANDER ISRAEL POVEDA GARCES](#) - Monday, 29 de May de 2023, 21:10

La distancia y la similitud son dos elementos vitales para la ciencia de datos, ya que nos permite efectuar análisis y mediciones dentro de los casos de estudios. Estos nos brindan medidas cuantitativas sobre dos objetos, logrando así determinar mediciones exactas. Por un lado, la **distancia** es una variable numérica que nos permite medir de forma cuantitativa cuán diferentes o separados se encuentran los objetos. Dependiendo del contexto o los datos analizados, puede tener diferentes formas de medir. Por otro lado, la **similitud** es una variable cuantitativa que mide la similitud o relación entre dos objetos, asimismo, dependiendo de sus datos y el contexto, puede medirse de diferentes formas. Esta variable es empleada para clasificar, identificar y agrupar valores por categoría.

Podemos ver presente esas dos variables en casos de estudios dentro de múltiples áreas como la ingeniería, la ecología, análisis de datos y estadística. Dentro de la ingeniería en sistemas podemos verla presente al momento de analizar grandes bases de datos en donde exista información relacionada a patrones, tendencias y consumos.

Como caso de estudio para aplicar estas variables, tenemos: **Crecimiento de la población (% Anual)**

<https://datos.bancomundial.org/indicador/SP.POP.GROW>

[Enlace permanente](#)

[Mostrar mensaje anterior](#)

Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [PAUL ALEXANDER GUARANDA MERO](#) - Monday, 29 de May de 2023, 21:31

La distancia y la similitud para comparar y medir la relación entre diferentes objetos o puntos de datos. Estos conceptos nos permiten cuantificar la proximidad o la diferencia entre ellos.

La distancia se refiere a una medida cuantitativa de la separación entre dos objetos o puntos en un espacio multidimensional. En el análisis de datos, se emplean diversas métricas de distancia comunes, la similitud se refiere a una medida que evalúa qué tan similares son dos objetos o puntos de datos entre sí. A diferencia de la distancia, la similitud busca medir la cercanía o semejanza entre ellos.

Mi base de análisis será los accidentes automovilísticos ocurridos en Ecuador

<https://www.ant.gob.ec/historico-estadisticas-siniestros-de-transito/#>

[Enlace permanente](#)

[Mostrar mensaje anterior](#)

Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [JOSELYN TATIANA RIVERA BAQUE](#) - Monday, 29 de May de 2023, 21:21

La distancia se refiere a la separación o espacio entre dos puntos, objetos o ubicaciones. Es una medida de qué tan separados o cerca están dos elementos en el espacio. La distancia se puede medir en varias unidades, como metros, kilómetros, millas, pulgadas, según el sistema de unidades utilizado.

Existen diferentes tipos de distancia, como la distancia euclídea, que se utiliza en geometría para medir la línea recta entre dos puntos en un plano o en un espacio tridimensional. También existe la distancia geográfica, que se refiere a la separación entre dos lugares en la superficie terrestre y se puede medir en kilómetros o millas.

Es importante señalar que la distancia puede ser relativa y variar según el marco de referencia utilizado. Por ejemplo, la distancia entre dos ciudades puede ser diferente si se mide en línea recta (distancia euclídea) o si se tiene en cuenta la red de carreteras y las rutas disponibles.

En resumen, la distancia es una medida de separación o espacio entre dos puntos y puede expresarse en diferentes unidades según el contexto.

La semejanza se refiere a la semejanza, semejanza o grado de semejanza entre dos o más cosas. Es un concepto utilizado para comparar y determinar hasta qué punto dos objetos, conceptos o entidades comparten características, características o rasgos comunes.

En varios campos, como las matemáticas, la estadística, la informática y la lingüística, la similitud a menudo se cuantifica mediante métodos matemáticos o computacionales. Estos métodos implican comparar atributos, propiedades o patrones de los objetos que

se analizan para evaluar su similitud.

Existen diferentes formas de medir la similitud, según el contexto y la naturaleza de los objetos que se comparan. Por ejemplo, en el análisis de texto, la similitud del coseno o la similitud de Jaccard se puede utilizar para determinar la similitud entre dos documentos en función de la superposición de sus frecuencias de palabras o conjuntos de palabras. En el procesamiento de imágenes, la similitud se puede evaluar utilizando técnicas como el índice de similitud estructural (SSIM) o el error cuadrático medio (MSE).

Es importante tener en cuenta que la similitud es un concepto subjetivo y se puede definir y medir de manera diferente según el contexto y los criterios específicos utilizados para la comparación.

Mi tema a analizar es: Conjunto de datos del censo de población del estado de Maharashtra, India.

Base de datos:<https://www.kaggle.com/datasets/bilalwaseer/microsoft-stocks-from-1986-to-2023>

[Enlace permanente](#)

[Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD

de [JOHNNY JOSE CRUZ BAQUERIZO](#) - Monday, 29 de May de 2023, 21:25

La similitud es un concepto amplio y se aplica en diversos campos del conocimiento. Aquí hay algunos puntos clave sobre la similitud:

1. Comparación de características: La similitud implica la comparación de características, atributos o cualidades de dos o más objetos, entidades o conceptos. Se evalúa la semejanza o igualdad en términos de estas características específicas.
2. Medición de similitud: La similitud puede ser medida de diferentes maneras, dependiendo del dominio o contexto en el que se esté evaluando. En algunos casos, se utilizan medidas cuantitativas, como la distancia euclídea o el coeficiente de similitud del coseno. En otros casos, se emplean métodos cualitativos o comparativos, basados en la observación de características comunes.
3. Contexto específico: La similitud es relativa al contexto en el que se aplica. Dos objetos pueden ser similares en un aspecto y diferentes en otro. Por lo tanto, la evaluación de similitud debe considerar el marco de referencia y los criterios específicos utilizados para la comparación.
4. Diversidad de aplicaciones: La similitud se utiliza en una amplia variedad de disciplinas. Por ejemplo, en la lingüística, se estudia la similitud entre idiomas o entre palabras. En la informática, se utiliza en algoritmos de clasificación, búsqueda de información y reconocimiento de patrones. También se aplica en ciencias sociales, biología, química, psicología y muchas otras áreas.
5. Técnicas de análisis: Para evaluar la similitud, se emplean diversas técnicas y herramientas, como algoritmos de aprendizaje automático, análisis de datos, técnicas estadísticas, minería de datos y procesamiento del lenguaje natural. Estas técnicas ayudan a identificar patrones, comparar características y cuantificar la similitud en conjuntos de datos o muestras.

En resumen, la similitud implica la comparación y evaluación de características comunes entre objetos, entidades o conceptos, y su medida puede variar según el contexto y los métodos utilizados.

La distancia se refiere a la separación física o espacial entre dos puntos, objetos, lugares o eventos. Es una medida cuantitativa que indica el espacio que los separa.

La distancia puede ser medida de diversas formas, dependiendo del contexto y del sistema de unidades utilizado. Algunas unidades comunes de distancia incluyen metros, kilómetros, millas, pulgadas, pies, entre otras. La elección de la unidad de medida depende de la escala y la precisión requerida en la medición.

Es importante tener en cuenta que la distancia puede ser medida en línea recta (distancia euclídea) o siguiendo una ruta específica (como la distancia en una carretera). Además, la distancia puede ser bidimensional (en un plano) o tridimensional (en el espacio).

La distancia es una propiedad fundamental en muchas disciplinas y áreas de estudio. Por ejemplo, en la física, la distancia es un factor crucial en el cálculo de velocidad, tiempo y aceleración. En la geografía, la distancia es esencial para describir la separación entre lugares y para la navegación. En la informática y las redes, la distancia se utiliza para determinar la eficiencia y la latencia en la transmisión de datos.

En resumen, la distancia se refiere a la separación física o espacial entre dos puntos, y es una medida fundamental en varios campos del conocimiento.

<https://www.kaggle.com/datasets/rajkumarpandey02/vehicle-fuel-economy-data-us>

[Enlace permanente](#)

[Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
 de [ANTHONY DOUGLAS FLORES TOMALA](#) - Monday, 29 de May de 2023, 21:34

Tanto la similitud y la distancia son utilizadas para comparar y medir la relación entre los diferentes objetos o conjuntos de datos utilizadas en una variedad de aplicaciones por ejemplo la clasificación, agrupación, recuperación de información y la recomendación de datos.

La similitud indica qué tan parecidos o cercanos son dos objetos o conjuntos de datos mientras que la distancia es una medida que indica la separación o diferencia entre dos objetos o conjuntos de datos, son herramientas esenciales en la ciencia de datos para analizar las relaciones entre los datos.

Tema: Inteligencia artificial para Power BI: Realizar un recorrido

[Ejemplo de inteligencia artificial para Power BI: Realizar un recorrido - Power BI | Microsoft Learn](#)

[Enlace permanente](#) [Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
 de [JAIME ANDRES VELEZ VERA](#) - Monday, 29 de May de 2023, 21:53

¿Qué es la distancia?

La palabra distancia proviene del latín distantia y significa 'aquellos que están establecidos lejos'.

La distancia se refiere a la longitud de la trayectoria que se recorre. Esto quiere decir que partiendo de un punto A hasta un punto B existe una longitud que une estos dos puntos, a esta longitud de la trayectoria se le denomina distancia.

¿Qué es el desplazamiento?

La palabra desplazamiento se refiere al movimiento que se realiza desde un lugar hasta otro. Cuando algo se mueve desde una posición inicial hasta una posición final se dice que hubo un desplazamiento.

La distancia y la dirección de la posición final con relación a la posición inicial es lo que nos indica cuál fue la longitud del desplazamiento.

Cuando hablamos de distancia nos referimos a la longitud entre dos puntos cualesquiera en el espacio, mientras que el desplazamiento se refiere a la distancia desde una posición inicial hasta una posición final independientemente del recorrido.
[link:https://www.diferenciador.com/distancia-y-desplazamiento/](https://www.diferenciador.com/distancia-y-desplazamiento/)

[Enlace permanente](#) [Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
 de [JULIAN STEVEN LAVAYEN GOMEZ](#) - Monday, 29 de May de 2023, 22:03

La distancia y similitud son herramientas utilizadas en la ciencia de datos para medir las relaciones entre relaciones de los elementos analizados, para identificar patrones de comportamiento que permitan dimensionar los resultados del estudio realizado.

Ejemplo: En procesos productivos, la distancia y similitud se puede emplear para medir la relación entre los proveedores de materia prima y el producto final obtenido, la similitudes de la materia prima de cada proveedor, y la distancia entre los resultados obtenidos con la materia prima de cada proveedor.

Datos para análisis:

Promedio de Graduados.

<https://www.kaggle.com/datasets/rkiattisak/graduation-rate>

[Enlace permanente](#) [Mostrar mensaje anterior](#)



Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD
 de [BETZABETH MADELINE MUÑOZ VILLEGAS](#) - Monday, 29 de May de 2023, 23:15

La distancia y la similitud son conceptos utilizados para comparar y agrupar elementos similares, identificar patrones y realizar análisis de similitud o clasificación.

La distancia se refiere a la medida de diferencia entre dos elementos, mientras que la similitud se relaciona con la comparación y la relación de semejanza entre dos elementos. Por ejemplo, la similitud entre conjuntos de elementos, como, por ejemplo, la similitud entre conjuntos de palabras clave en un análisis de texto.

Tema: Brecha salarial entre géneros**Data:** [Brecha salarial según sexo, antigüedad, formación o categoría profesional, según Adecco e Infoempleo \(epdata.es\)](#)

Calificación máxima:10 (1)

[Enlace permanente](#)[Mostrar mensaje anterior](#)**Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD**de [FRANCISCO JAVIER VALVERDE CHAGUAY](#) - Monday, 29 de May de 2023, 23:21

En el campo de la ciencia de datos, se utilizan la similitud y la distancia para visualizar patrones de relaciones entre variables. La distancia mide la diferencia entre dos objetos o puntos, mientras que la similitud mide la proximidad o similitud entre ellos. Ambos conceptos son medidas cuantitativas y se usan en técnicas de análisis de datos. La distancia permite determinar la cercanía o lejanía entre observaciones, pero puede ser afectada por valores atípicos. Por otro lado, las similitudes se utilizan para agrupar observaciones similares, pero pueden no reflejar correctamente las diferencias si las variables tienen diferentes escalas o distribuciones. En conjunto, estas herramientas son valiosas para comprender y analizar datos en diversas áreas de la ingeniería.

Tema: Acciones de Microsoft de 1986 a 2023

Link: <https://www.kaggle.com/datasets/bilalwaseer/microsoft-stocks-from-1986-to-2023>[Enlace permanente](#)[Mostrar mensaje anterior](#)**Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD**de [ELVIN DANIEL IBARRA RONQUILLO](#) - Monday, 29 de May de 2023, 23:32

La distancia se refiere a la medida de separación o diferencia entre dos objetos, puntos o elementos en un espacio determinado. Existen diferentes tipos de distancias, y la elección de la medida adecuada depende del contexto y del tipo de datos que se esté analizando.

Distancias más comunes:

- **Distancia euclíadiana**
- **Distancia de Manhattan**
- **Distancia de Minkowski**.
- **Distancia de Hamming**

La similitud, por otro lado, se refiere a una medida de cuán parecidos son dos objetos o elementos. La similitud puede ser cuantificada mediante diferentes métricas o coeficientes de similitud, dependiendo del tipo de datos y del contexto de análisis.

Similitudes más comunes:

- **Coeficiente de correlación**
- **Coeficiente de Jaccard**
- **Coeficiente de coseno**
- **Distancia de Levenshtein**

Base de dato:[Students Exam Scores: Extended Dataset | Kaggle](#)[Enlace permanente](#)[Mostrar mensaje anterior](#)**Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD**de [ARTURO ENRIQUE ROSERO ESPINOZA](#) - Monday, 19 de June de 2023, 19:50

Los científicos de datos usan análisis de distancia y similitud para visualizar patrones y relaciones entre variables. La similitud mide su proximidad o semejanza, la distancia mide la diferencia entre dos objetos o puntos. Para explorar y visualizar patrones y relaciones entre objetos, las técnicas estadísticas y de análisis de datos utilizan ambos conceptos.

Las mediciones de distancia son fáciles de comprender y calcular, pero es posible que no capturen con precisión las similitudes entre las observaciones realizadas utilizando varias escalas o unidades.

Las medidas de similitud son útiles para determinar qué tan estrechamente relacionadas están las observaciones y pueden revelar información sobre la dirección de las relaciones, pero pueden ser más difíciles de calcular y comprender.

[Enlace permanente](#)[Mostrar mensaje anterior](#)**Re: FORO 2-1: RESUMEN SOBRE CONCEPTOS DE DISTANCIA Y SIMILITUD**de [GENESIS LISSET SANCHEZ MORAN](#) - Monday, 19 de June de 2023, 20:52

La distancia y la similitud son dos conceptos básicos en el análisis de datos y las estadísticas. La distancia es una medida cuantitativa de la diferencia entre dos objetos o puntos en una habitación. La similitud, por otro lado, es una medida cuantitativa de la proximidad o similitud entre dos objetos o puntos. La similitud se refiere al grado de dos objetos o puntos similares entre sí. La similitud se puede medir de varias maneras, dependiendo del tipo de datos analizados.

◀ Investigar y documentar sobre los usos de la aplicación Power BI para la Ciencia de Datos

Ir a...

FORO 2-2: RESUMEN SOBRE TECNICAS DE SIMILITUD Y DISTANCIA ►

Usted se ha identificado como BETZABETH MADELINE MUÑOZ VILLEGAS (Cerrar sesión)

Reiniciar tour para usuario en esta página

15150059_15150091_1515811

Español - Internacional (es)

English (en)

Español - Internacional (es)

Resumen de retención de datos

Descargar la app para dispositivos móviles