**Ficha estadística**

**Introducción**

La introducción de un resumen estadístico en el informe de ventas de Adventure Works encendió una chispa de entusiasmo. El equipo directivo reconoció la importancia de aplicar estadísticas a los conjuntos de datos de Adventure Work y acordó explorar todas las capacidades de las herramientas estadísticas. Mientras se prepara para adaptarse a estas necesidades, decide explorar las capacidades estadísticas de Microsoft Power BI. Power BI libera todo el potencial de la estadística, proporcionando una variada caja de herramientas de funciones estadísticas, elaboradas para satisfacer las necesidades específicas de las empresas de hoy en día.

El objetivo de esta lectura es proporcionarle una lista de las funciones estadísticas de Power BI para que comprenda mejor la funcionalidad de cada función estadística, a continuación se proporciona un conjunto de datos de muestra. Al observar los resultados de cada función con los datos de muestra, podrá comprender la aplicación práctica de cada función DAX y cómo pueden utilizarse en escenarios del mundo real.

Conjunto de datos de ventas de muestra:

| **Precio\_artículo** | **Precio\_artículo\_con Texto** | **CantidadVendida** |
| --- | --- | --- |
| 120 | 120 | 5 |
| 140 | Verdadero | 7 |
| 180 | Texto aleatorio | 3 |
| 130 | 130 | 6 |
| 200 | 180 | 4 |
| 240 | Falso | 9 |
| 160 | 280 | 2 |
| 190 |  | 8 |

Para cada función DAX vista en la siguiente lectura, encontrará su expresión, una definición clara y un ejemplo de cómo funciona con el conjunto de datos de muestra proporcionado.

**Funciones de media y mediana**

Las funciones de**media** y **mediana** son esenciales en casi cualquier conjunto de datos que contenga columnas numéricas. Por ejemplo, en una tienda en línea, puede utilizar estas funciones para evaluar los valores medios y medianos de los pedidos, proporcionando información sobre los hábitos de gasto de los clientes. Del mismo modo, como profesor que evalúa las puntuaciones de los alumnos en los exámenes, el cálculo de las puntuaciones media y mediana ofrece una visión rápida del rendimiento de la clase a la vez que identifica los valores atípicos.

**Funciones de media**

**1. PROMEDIO(<columna>):** El valor medio de la columna dada se calcula sumando todos los valores y dividiéndolo por el número de puntos de datos.

Ejemplo: PROMEDIO **([PrecioArtículo])** Por ejemplo, si la columna **ItemPrice** contuviera 8 valores, la fórmula se calcularía como (120 + 140 + 180 + 130 + 200 + 240 + 160 + 190) / 8 = 170

**2. AVERAGEA(<columna>)**: Se calcula el valor medio de la columna dada, manejando tipos de datos no numéricos basándose en reglas específicas.

**Las reglas son las siguientes:** **- Los valores que se evalúan como TRUE cuentan como 1.**

- Los valores que se evalúan como FALSE cuentan como 0 (cero).

- Los valores que contienen texto no numérico cuentan como 0 (cero).

- El texto vacío ("") cuenta como 0 (cero).

Ejemplo: **AVERAGEA([PrecioArtículo\_conTexto])**

Por ejemplo, si la columna **Precio\_artículo\_conTexto** contuviera 8 valores, contando los valores no numéricos como contados como 0, la fórmula se calcularía como(120 + 1 + 0 + 130 + 180 + 0 + 280 + 1) / 8 = 89

**3. AVERAGEX(<tabla>, <expresión>)**: Se calcula la media de la expresión en la tabla especificada. Se evalúa la expresión para cada fila de la tabla y utiliza la media de los resultados.

Ejemplo: **AVERAGEX(Ventas, [PrecioArtículo] \* [CantidadVendida])** Por ejemplo, si la columna **PrecioArtículo** y la columna **CantidadVendida** contuvieran 8 valores cada una, la fórmula se calcularía como ((120 \* 5) + (140 \* 7) + (180 \* 3) + (130 \* 6) + (200 \* 4) + (240 \* 9) + (160 \* 2) + (190 \* 8)) / 8 = 962.50

**Función mediana**

**1. MEDIAN(<columna>)**: Devuelve el valor de la mediana de una columna numérica.

Ejemplo: MEDIAN **(Ventas[PrecioArtículo])** Por ejemplo, si la columna **PrecioArtículo** de la tabla **Ventas** contuviera 2 valores, la fórmula se calcularía como (160 + 180) / 2 = 170

**Funciones de varianza y desviación estándar**

Tanto la varianza como la desviación estándar son medidas de lo extendido o disperso que está un conjunto de puntos de datos, por lo que desempeñan un papel fundamental cuando es necesario medir la variabilidad de los datos. En el contexto de la gestión de carteras financieras, calcular la varianza y la desviación típica de los rendimientos de las acciones ayuda a calibrar el riesgo de las inversiones. En la fabricación, la supervisión del espesor de la pintura en una planta de automóviles implica determinar la varianza y la desviación estándar de las mediciones para poner de manifiesto las incoherencias del proceso. El cálculo de estas dos funciones difiere ligeramente dependiendo de si el conjunto de datos representa toda la población de puntos de datos o sólo una muestra de ella. En DAX, esta diferenciación se denota mediante una P para población y una S, para una muestra en sus fórmulas. Utilizar la fórmula de varianza o desviación típica adecuada es esencial para garantizar un análisis preciso y significativo basado en la naturaleza de los datos disponibles.

**Varianza y desviación típica utilizando una muestra de población**

Supongamos que la columna sobre la que se calculará la varianza/desviación se refiere a una población muestral:

**1. VAR.S(<columna>)**: Devuelve la varianza de una columna que contiene una población muestral.

Ejemplo: VAR.S **(Ventas[PrecioArtículo])** = 1.629

**2. STDEV.S(<columna>)**: Devuelve la desviación estándar de una columna que contiene una población de muestra.

Ejemplo: **STDEV.S(Ventas[PrecioArtículo])**= 40

**Varianza y desviación típica utilizando toda la población**

Suponiendo que la columna sobre la que se calculará la varianza/desviación contiene toda la población de puntos de datos:

**1. VAR.P(<columna>)**: Devuelve la varianza de una columna, suponiendo que la columna se refiere a toda la población.

Ejemplo: VAR.P **(Ventas[PrecioArtículo])** = 1.425

**2. STDEV.P(<columna>)**: Devuelve la desviación estándar de una columna que contiene toda una población.

Ejemplo: **STDEV.P(Ventas[PrecioArtículo])** = 38

**Funciones de recuento, mín y máx**

Contar, mín. y máx. son funciones versátiles con diversas aplicaciones en cualquier proyecto de análisis de datos. Por ejemplo, en el servicio de atención al cliente, contar los tickets resueltos, hallar los tiempos mínimos de resolución e identificar los índices máximos de satisfacción del cliente, muestran los logros individuales de los agentes. Del mismo modo, en la gestión de inventarios minoristas, contar los artículos, hallar los niveles mínimos de existencias e identificar los volúmenes máximos de ventas contribuyen a optimizar los niveles de inventario.

**Funciones de recuento**

**1. COUNT(<columna>)**: Cuenta el número total de filas, conteniendo cualquier valor o ninguno.

Ejemplo: **COUNT([PrecioArtículo\_conTexto])** = 8

**2. COUNTBLANK(<columna>)**: Cuenta el número de filas en blanco o vacías de una columna.

Ejemplo: COUNTBLANK([PrecioArtículo\_conTexto **])**= 1

**3. DISTINCTCOUNT(<columna>)**: Cuenta el número de valores distintos en la columna.

Ejemplo: DISTINCTCOUNT([Precio\_artículo\_conTexto **])** = 8

**Función Max**

**1. MAX(<columna>)**:Devuelve el mayor valor numérico o la mayor cadena de una columna. Ignora los valores lógicos.

Ejemplo **MAX([Precio\_artículo\_conTexto])** = "TRUE"

**Nota:** Como la función MAX ignora los valores lógicos, TRUE y FALSE en la tabla se interpretan como textos. Los textos se consideran "más grandes" que los números en DAX alfabéticamente, por lo que el valor más grande de la columna es TRUE, ya que es el último texto en orden alfabético.

**Función Mín**

**1. MIN(<columna)**: Devuelve el valor numérico más pequeño o la cadena más corta de una columna. Ignora los valores lógicos.

Ejemplo MIN**([PrecioArtículo\_conTexto])** = ""

**Nota:** Como la función MIN también ignora los valores lógicos, al igual que la función MAX, la secuencia de valores en blanco, números y textos se sigue en orden ascendente. Un valor en blanco en la columna de muestra se considera como el valor mínimo.

**Funciones de combinación y permutación**

Las funciones de combinación y permutación en estadística calculan el número de combinaciones para un número determinado de elementos. La diferencia entre las dos funciones es que en una función de combinación, el orden de salida de los elementos no es importante, mientras que en una función de permutación, el orden de salida de los elementos es significativo. Para poner un ejemplo sencillo, calculemos las probabilidades de ganar un billete de lotería, en el que hay que acertar 6 números de 49, utilizando estas dos funciones en Power BI.

**Función de combinación**

**1. COMBIN(n,k**): Devuelve el número de formas en que se pueden seleccionar **k** elementos de **n**, donde el orden no importa.

Ejemplo: **COMBIN(49,6)** = 13.983.816

**Función de permutación**

**1. PERMUT(n,k):** Devuelve el número de formas en que se pueden seleccionar **k elementos** de **n**, donde el orden sí importa.

Ejemplo: **PERMUT(49,6)** = 10.068.347.520

**Conclusión**

Esta hoja de trucos estadísticos ha cubierto las funciones estadísticas más utilizadas en Power BI, aprovechando las capacidades del lenguaje DAX. Proporcionó una amplia gama de funciones, desde cálculos sencillos como medias y medianas hasta otros más complejos como la desviación estándar, y funciones de probabilidad, calculando las probabilidades de ganar un billete de lotería. Utilizando los ejemplos prácticos con el conjunto de datos de muestra proporcionado en esta lectura, ahora es capaz de identificar, explicar y utilizar las diferentes funciones estadísticas disponibles en DAX.