

LECCIÓN 12.6

Representación de datos con diagramas de caja



Normas educativas federales *Common Core*

S-ID.1 Representar datos con diagramas en la recta numérica real (diagrama de puntos, histogramas y diagramas de caja).

S-ID.2 Utilizar estadísticas adecuadas a la forma de distribución de datos para comparar el centro (mediana, media) y la distribución (rango intercuartil, desviación estándar) de dos o más conjuntos de datos diferentes.

S-ID.3 Interpretar las diferencias de forma, centro y distribución en el contexto de los conjuntos de datos, teniendo en cuenta los posibles efectos de los puntos de valores extremos (valores atípicos).

Prácticas de matemáticas 2, 4, 5

ENFOQUE DE LA PROFESIÓN: gerente de sistemas informáticos y de información

Los gerentes de sistemas informáticos y de información, comúnmente conocidos como gerentes de tecnología de la información (TI), planifican, coordinan y dirigen las actividades relacionadas con la computación en una organización. Aplican su experiencia en ciencias informáticas y matemáticas para resolver problemas, que están determinados por el negocio de las organizaciones para las que trabajan.

- Hable sobre la tecnología de la información con los estudiantes leyendo juntos Enfoque de la profesión.
- Busque institutos y universidades locales con programas en ciencias informáticas y tecnología de la información para compartir con los estudiantes.
- Investigue compañías locales que emplean gerentes de sistemas informáticos y de información y pregúnteles qué tipo de trabajo hacen en las compañías.

Video: gerentes de sistemas informáticos y de información

Muestre este video a los estudiantes, el cual describe los tipos de proyectos en los que gerentes de sistemas informáticos y de información podrían trabajar.

Objetivo de la lección

En esta lección, verá cómo un gerente de sistemas informáticos y de información puede usar diagramas de caja para analizar e interpretar datos.

1 Una muestra de la carrera: cómo crear diagramas de caja

A los empleados de la compañía les preocupa que los mensajes de correo electrónico a un cliente importante se entreguen con demasiada lentitud. El gerente de proyectos de informática comienza su investigación contando el número de mensajes de correo electrónico que 15 empleados enviaron a este cliente durante el último mes. Las cifras se muestran en la lista.

27, 32, 17, 15, 9, 36, 23, 11, 8, 15, 20, 27, 34, 22, 28

Dibuja un diagrama de caja con los datos. Luego aplica la mediana, el rango y el rango intercuartil para interpretar los datos.



Preguntas orientadoras

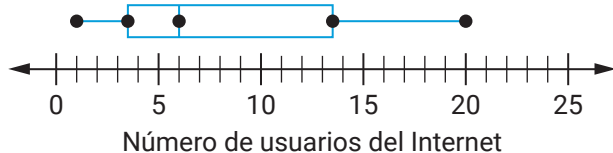
- En el paso 1, ¿por qué es importante enumerar los datos en orden de menor a mayor?
- En el paso 1, ¿cuántos valores de datos hay? ¿Cómo les ayuda este número a determinar la mediana?
- Algunos de los valores de datos aparecen más de una vez en la lista. En el paso 1, ¿es necesario repetir estos valores en la lista ordenada que generaron? ¿Por qué?

DISCERNIMIENTO: SUPLEMENTO Recuerde a los estudiantes que en este problema están usando un conjunto de 15 valores de datos para crear un diagrama de caja. Pregúnteles si es posible trabajar al revés, comenzando con un diagrama de caja para determinar el conjunto de datos. Por ejemplo, si se les diera el diagrama de caja para este ejemplo, ¿podrían generar la lista de valores que representa? Explique que la respuesta es no. Un diagrama de caja podría representar varios conjuntos de datos diferentes. Como analogía, explique que un diagrama de caja es como el resumen de una novela o una historia. Es posible escribir un resumen útil después de leer la historia, pero muchos detalles de una historia se omiten en el resumen.

En el trabajo: aplicar la creación de diagramas de caja

Respuestas

1a.



1b. rango: 19; rango intercuartil: 10

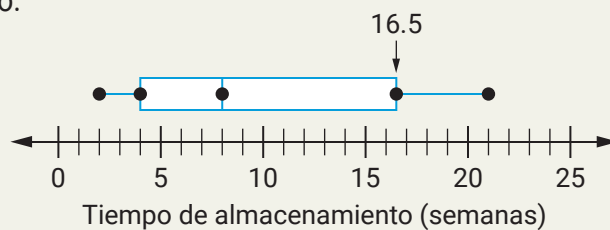
1c. La distribución está sesgada a la derecha, lo que significa que los datos superiores a la mediana están más distribuidos, mientras que los datos inferiores a la mediana están agrupados.

Utilice estas preguntas para verificar la comprensión de los estudiantes.

- En 1a, ¿cómo afectó el número par la manera en que calcularon la mediana?
- En 1b, ¿cómo determinaron el rango? ¿Cómo determinaron el rango intercuartil?
- ¿Qué datos adicionales ayudarían a Harlan a evaluar el uso del Internet por parte de los clientes de los restaurantes?

2 Una muestra de la carrera: cómo interpretar los diagramas de caja

Jeanine es la directora de informática de una pequeña empresa. Le preocupa que los empleados de la empresa no estén eliminando los archivos grandes del servidor cuando ya no se necesiten. Ella reúne los datos sobre los archivos más grandes que están en el servidor. El diagrama de caja resume el tiempo de almacenamiento en semanas de cada archivo.



Utiliza el resumen de cinco números, el rango y el IQR para describir los datos que se muestran en el diagrama de caja.

Preguntas orientadoras

- En el Paso 1, ¿cómo determinan la mediana de los datos?
- ¿Cómo indica el diagrama de caja las distribuciones de los valores de datos inferiores a la mediana y los valores de datos superiores a la mediana?
- El diagrama de caja indica que los valores de datos incluyen un tiempo mínimo de 2 semanas y un tiempo máximo de 21 semanas. ¿El diagrama indica otros valores específicos en el conjunto de datos?

SUPLEMENTO Pida a los estudiantes que propongan un conjunto de valores de datos que estarían representados por el diagrama de caja que se indica para este problema. Los estudiantes deben comenzar con la inclusión de cinco valores en el resumen de cinco números de los datos.

En el trabajo: aplicar la interpretación de los diagramas de caja

Respuestas

- 2a. mínimo: 1 hora; Q_1 : 9 horas; mediana: 11.5 horas; Q_3 : 20 horas; máximo: 23 horas; rango: 22 horas; IQR: 11 horas
- 2b. Los datos se distribuyen de una manera algo uniforme, aunque los valores entre el primer cuartil y la mediana están agrupados. Los valores cercanos al mínimo se distribuyen más ampliamente que los que están cerca del máximo.

Utilice estas preguntas para verificar la comprensión de los estudiantes.

- En la 2a, ¿cómo utilizaron el diagrama de caja para determinar el resumen de cinco números?
- En la 2a, ¿cómo determinaron el rango? ¿Cómo determinaron el IQR?
- En la 2b, ¿de qué manera les ayudó la información del resumen de cinco números a describir los datos?

3 Una muestra de la carrera: cómo comparar los diagramas de caja

Una empresa grande proporciona teléfonos móviles a todos sus empleados. Naomi es una de las gerentes de informática de la empresa. Ella le pide a un proveedor de servicios telefónicos sobre el uso de estos teléfonos el número de horas de uso del teléfono durante un mes reciente para los representantes de ventas en la región oeste y en la región este.

Región oeste: 10, 11, 9, 34, 36, 25, 28, 21, 10, 22, 5, 8, 32

Región este: 22, 16, 10, 20, 10, 18, 14, 8, 10, 21, 22, 11, 12

Dibuja un diagrama de caja doble para comparar los conjuntos de datos. Utiliza los diagramas de caja para comparar las horas de uso del teléfono entre los representantes de ventas de las dos regiones.



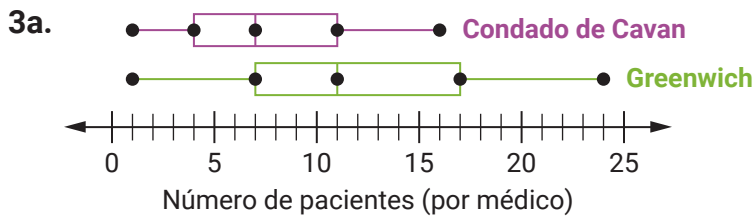
Preguntas orientadoras

- ¿Es razonable utilizar la misma escala en ambos diagramas de caja?
- ¿Creen que es más prudente completar un diagrama de caja y luego completar el otro, o deberían completar los dos diagramas de caja juntos llevando a cabo el mismo paso en cada conjunto de datos? Den una explicación.
- ¿Por qué es importante etiquetar correctamente los diagramas de caja?
- ¿Qué pasa si algunos de los representantes de ventas trabajan en las dos regiones y sus datos aparecen en ambas listas? ¿Cómo afectaría esto al análisis de los datos?

DISCERNIMIENTO: COMPLEMENTO Hable sobre cómo los teléfonos celulares, las computadoras portátiles y otros dispositivos electrónicos se han vuelto esenciales en muchas empresas y las empresas a menudo compran estos dispositivos para que sus empleados los utilicen. Cuando los empleados trabajan fuera de una oficina central, los gerentes de informática generalmente los atienden por correo electrónico o por teléfono. A veces, el empleado y el gerente de informática se envían equipos entre ellos a través del correo postal o de entrega de paquetes.

En el trabajo: aplicar la comparación de los diagramas de caja

Respuestas



- 3b. Los datos de las dos oficinas tienen el mismo mínimo, pero los cuartiles, la mediana y el máximo son valores mayores para la oficina de Greenwich que para la del condado de Cavan. Los datos también se distribuyen más ampliamente y en un rango mayor en la oficina de Greenwich.

Utilice estas preguntas para verificar la comprensión de los estudiantes.

- ¿En qué oficina aceptaron el sitio web más pacientes? ¿Cómo pueden saberlo?
- Para cada oficina, ¿aproximadamente qué fracción de los médicos tiene entre 7 y 11 pacientes que aceptaron el sitio web?
- ¿Qué factores podrían explicar la diferencia entre los dos conjuntos de datos? ¿Qué información adicional sobre los médicos o sus pacientes les ayudaría a evaluar sus ideas?

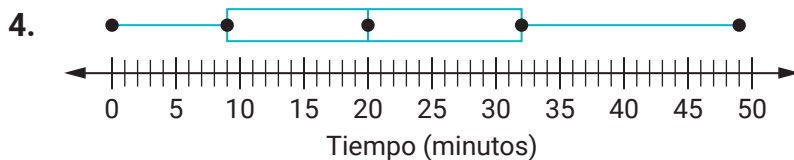
Enfoque de la profesión: práctica

Pasos para la solución de los ejercicios 4 a 7

Estos pasos ayudarán a guiar a los estudiantes a resolver estos ejercicios de práctica.

Ejercicio 4

Respuesta



Diseña un plan

Plan posible:

Paso 1: organizar los datos en orden ascendente.

Paso 2: determinar el resumen de cinco números del conjunto de datos.

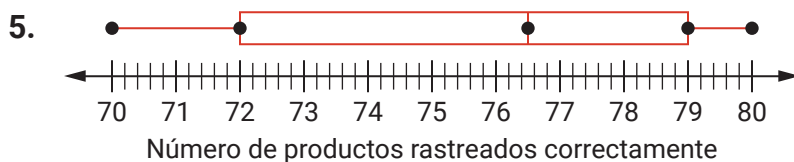
Paso 3: dibujar el diagrama de caja.

Pasos de la solución

- Ordenar los datos de menor a mayor. (0, 4, 8, 10, 15, 20, 20, 24, 30, 34, 38, 49)
- Identificar los valores mínimos y máximos. (mínimo: 0; máximo: 49)
- Determinar la mediana. Debido a que la cantidad de los valores de datos es un número par, la mediana es el promedio de los dos valores que están en medio. (mediana: 20)
- Determinar el primer y tercer cuartiles, que son las medianas de las mitades inferior y superior de los datos. (primer cuartil: 9; tercer cuartil: 32)
- Elegir una escala adecuada y dibujar el resumen de cinco números del conjunto de datos.
- Conectar los puntos con líneas y un cuadro para formar el diagrama de caja.

Ejercicio 5

Respuesta



- Ordenar los datos de menor a mayor. (70, 71, 72, 75, 76, 77, 79, 79, 80, 80)
- Identificar los valores mínimos y máximos. (mínimo: 70; máximo: 80)
- Determinar la mediana. Debido a que la cantidad de los valores de datos es un número par, la mediana es el promedio de los dos valores que están en medio. (mediana: 76.5)
- Determinar el primer y tercer cuartiles, que son las medianas de las mitades inferior y superior de los datos. (primer cuartil: 72; tercer cuartil: 79)
- Elegir una escala adecuada y dibujar el resumen de cinco números del conjunto de datos.
- Conectar los puntos con líneas y un cuadro para formar el diagrama de caja.

Ejercicio 6

Respuestas

- 6a.** mínimo: 0; primer cuartil: 2.5; mediana: 7.5; tercer cuartil: 9; máximo: 12.5
- 6b.** El rango es 12.5. El rango indica que la diferencia entre cualquiera de dos computadoras en sus años de uso restantes es, como máximo, 12.5 años.
- 6c.** El rango intercuartil es 6.5. El rango intercuartil indica que para la mitad de las computadoras, la distribución de sus años de uso restantes es, como máximo, 6.5 años, y estos valores se acercan a la mediana.
- Identificar los cinco puntos que se describen en el diagrama de caja. (0, 2.5, 7.5, 9, 12.5)
 - Utilizar los cinco puntos para indicar el resumen de cinco números del conjunto de datos.
 - Encontrar el rango restando el mínimo al máximo. ($12.5 - 0 = 12.5$)
 - Encontrar el rango intercuartil restando el primer cuartil al tercer cuartil. ($9 - 2.5 = 6.5$)
 - Interpretar los significados del rango y el rango intercuartílico en el contexto del problema.

Ejercicio 7

Respuestas

- 7a. para Software A: mínimo: 35; primer cuartil: 38; mediana: 40; tercer cuartil: 45; máximo: 50 para Software B: mínimo: 25; primer cuartil: 35; mediana: 40; tercer cuartil: 45; máximo: 65
- 7b. para Software A: rango: 15; rango intercuartil: 7 para Software B: rango: 40; rango intercuartil: 10
- 7c. Los dos conjuntos de datos tienen la misma mediana, pero los datos del software B se distribuyen en un rango mayor. Los datos sugieren que los miembros del personal podrían llegar a ser competentes con el software B y producir documentos con relativa rapidez o que podrían tener problemas con el software y trabajar más lentamente. Los miembros del personal tuvieron menos variación en el tiempo que necesitaban para producir un documento con el software A.
- Identificar los cinco puntos para el software A que se describen en el diagrama de caja. (35, 38, 40, 45, 50)
 - Utilizar los cinco puntos para indicar el resumen de cinco números del conjunto de datos para el software A. (mínimo: 35, primer cuartil: 38, mediana: 40, tercer cuartil: 45, máximo: 50)
 - Identificar los cinco puntos para el software B que se describen en el diagrama de caja. (25, 35, 40, 45, 65)
 - Utilizar los cinco puntos para indicar el resumen de cinco números del conjunto de datos para el software B. (mínimo: 25, primer cuartil: 35, mediana: 40, tercer cuartil: 45, máximo: 65)
 - Encontrar el rango del software A restando el mínimo al máximo. ($50 - 35 = 15$)
 - Encontrar el rango intercuartil del software A restando el primer cuartil al tercer cuartil. ($45 - 38 = 7$)
 - Encontrar el rango del Software B restando el mínimo al máximo. ($65 - 25 = 40$)
 - Encontrar el rango intercuartil del software A restando el primer cuartil al tercer cuartil. ($45 - 35 = 10$)
 - Comparar la mediana y la distribución de los dos conjuntos de datos e interpretar los significados.

Enfoque de la profesión: verificación

Consejos para completar los ejercicios 8 a 13

Estos consejos ayudarán a los estudiantes a resolver estos ejercicios y elementos de evaluación similares.

Ejercicio 8

Respuesta

8. B

Consejo Los estudiantes pueden preguntarse cómo los valores repetidos en un conjunto de datos afectan el resumen de cinco números y el diagrama de caja. Para este problema, sugiera que los estudiantes reemplacen uno de cada valor repetido (5.5 y 6) con un número que sea un poco más pequeño o más grande. Luego, pídeles que busquen el resumen de cinco números de los datos modificados. Los estudiantes deben ver que los diagramas de caja son idénticos o solo un poco diferentes. Hable sobre cómo el modo, que es el valor más frecuente en un conjunto de datos, no se representa en un diagrama de caja.

Ejercicio 9

Respuesta

9. A: mínimo, B: primer cuartil, C: mediana, D: tercer cuartil, E: máximo

Consejo Recuerde a los estudiantes el significado de la palabra *cuartil*, que es una división en cuartos. Luego, pregunte a los estudiantes si pueden identificar el segundo cuartil en el conjunto de datos. Explique que el segundo cuartil es lo mismo que la mediana.

Ejercicio 10

Respuesta

10. C. ordenar los valores, b. de en medio, a. mediana, b. resta

Consejo Después de que los estudiantes completen este ejercicio, pídales que sigan los pasos del párrafo para practicar cómo encontrar un resumen de cinco números. Luego pídales que dibujen el diagrama de caja.

Ejercicio 11

Respuesta

11. Mediana: 6, primer cuartil: 3, rango intercuartil: 7, rango: 15

Consejo A medida que revisa este problema, señale que no indica el número de empleados en la empresa. Pregunte a los estudiantes si pueden determinar el número de empleados a partir del diagrama de caja. Hable sobre cómo los diagramas de caja resumen la distribución de datos, pero no muestran el número de valores.

Ejercicio 12

Respuesta

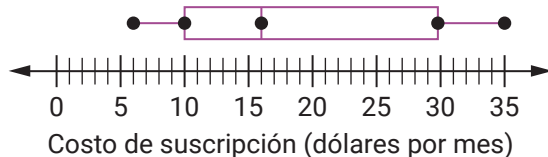
12. a, d

Consejo Hable sobre cómo los valores atípicos afectan los diagramas de caja. El efecto más obvio es una extensión de la línea horizontal que conecta el mínimo o el máximo con el resto del diagrama de caja. Otras partes del diagrama de caja pueden permanecer igual o cambiar solo un poco. En la situación descrita en la opción de respuesta e, la mediana y los cuartiles de los datos suburbanos pueden cambiar o no, mientras que el máximo seguirá siendo el mismo.

Ejercicio 13

Respuesta

13.



Consejo Recuerde a los estudiantes la importancia de elegir la escala correcta al dibujar cualquier gráfico, incluido un diagrama de caja. Analice cómo una escala demasiado pequeña puede hacer que un diagrama de caja sea demasiado grande, mientras que una escala demasiado grande puede reducir la apariencia de la caja y dificultar su interpretación.