***Guía de Ejercicios Prácticos Nº2***

Ejercicio Nº1: Considere las variables Edad y Estatura de la encuesta realizada y obtenga las medidas correspondientes mediante InfoStat para contestar las siguientes preguntas:

1. ¿cuál es la edad promedio de los estudiantes de Estadística I - 2022? ¿y la estatura promedio?
2. ¿qué edades tiene el 50% de los alumnos más jóvenes? ¿y de los más bajos?
3. ¿Cuál es la edad que más se repite en el curso? ¿y la estatura?
4. ¿Qué edades tiene el 20% de los alumnos con mayor edad? ¿qué estatura tienen el 65% de los alumnos?
5. ¿La distribución de las edades se distribuye en forma simétrica o asimétrica? ¿y el de las estaturas?
6. Construya un histograma y un gráfico de caja para los dos conjuntos de datos. ¿Hay algún valor atípico? ¿el gráfico de caja confirma su descripción acerca de la forma de la distribución?
7. ¿En cuál de las dos variables hay mayor dispersión de datos?

Ejercicio Nº2: Realizar un pequeño informe estadístico de los estudiantes que están cursando Estadística I – 2022 en cuanto a las variables Cantidad de Materias Aprobadas y Cantidad de Materias Regulares de la encuesta realizada.

Ejercicio Nº3: Un número significativo de empresas en bancarrota incurren en retrasos al publicar los reportes antes de la quiebra. Los tiempos de retrasos en años (también llamados períodos de demora) en estos reportes para 57 empresas quebradas se describen mediante el siguiente gráfico. Examine y conteste las siguientes preguntas:

1. Identifique unidad de análisis, variable y clasificación.
2. ¿Es cierto que el periodo de demora más frecuente de las empresas quebradas fue de 5 años? Justifique.
3. ¿Cuál es el tiempo promedio de retraso?
4. De los que menos años tardaron en presentar la quiebra ¿qué tiempo lo hicieron el 50% de las empresas?
5. De los que menos tardaron en presentar informe de quiebra ¿cuántos años tardó el 10% de las empresas?
6. De los que más años tardaron, ¿qué tiempo lo hicieron el 25% de las empresas?

Ejercicio Nº4: Dos formas que usan los alumnos para ir a la escuela diariamente son el transporte público y el automóvil de sus padres. A continuación, vemos unas muestras de tiempos de cada forma de transportarse. Las cifras son en minutos:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Transporte público | 25 | 28 | 29 | 29 | 32 | 32 | 33 | 34 | 37 | 41 |
| Automóvil | 29 | 30 | 31 | 31 | 32 | 32 | 33 | 33 | 34 | 35 |

1. Calcule la media y la mediana para conjunto de datos. Interprete cada una.
2. Determine la desviación estándar para cada conjunto de datos e interprételas.
3. Si tuvieras la posibilidad de recomendar a un alumno que generalmente llega tarde ¿Qué forma de transporte le recomendarías para llegar a la escuela? Justifique estadísticamente su respuesta.

Ejercicio Nº5: Se escogió una muestra de 714 conductores de ómnibus y se registró en la tabla siguiente el número de accidentes de tránsito que tuvieron durante los años 2006 y 2010.

|  |  |  |
| --- | --- | --- |
| **Número de accidentes** | **Cantidad de conductores** | En relación a esta distribución, complete incluyendo los cálculos necesarios:   1. El 40% de los conductores no supera los............................accidente/s. 2. El número promedio de accidentes en los cuatros años es de ....................... accidentes. 3. ¿A partir de qué número de accidentes se encuentra el 25% de los conductores con mayor número de accidentes? 4. ¿A partir de qué número de accidentes se encuentra el 10% de los conductores con mayor número de accidentes? 5. Obtenga el desvío estándar y el coeficiente de variación e interprételos. 6. Realice el diagrama de caja. ¿Qué puede decir acerca del conjunto de datos? |
| **0** | 114 |
| **1** | 157 |
| **2** | 158 |
| **3** | 115 |
| **4** | 78 |
| **5** | 44 |
| **6** | 21 |
| **7** | 7 |
| **8** | 6 |
| **9** | 10 |
| **10** | 3 |
| **11** | 1 |

Ejercicio Nº6: Suponga que un investigador desea estudiar como varía el peso, en kilogramos, de los paquetes de alimento para perros envasado por una fábrica de alimentos para animales domésticos. Para ello toma una muestra de 50 paquetes y registra los siguientes pesos agrupados.

|  |  |
| --- | --- |
| Intervalo | Frecuencia |
| 53 – 56 | 2 |
| 56 – 59 | 5 |
| 59 – 62 | 9 |
| 62 – 65 | 15 |
| 65 – 68 | 12 |
| 68 – 71 | 5 |
| 71 – 74 | 2 |

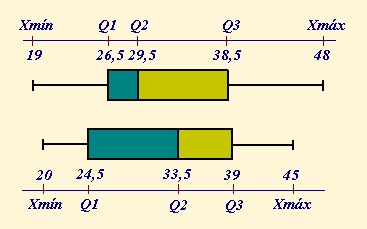
1. Obtenga la media, la moda y la mediana.
2. Realice un gráfico de caja.
3. Obtenga el coeficiente de variación y de asimetría. Interprételos.

Ejercicio Nº7: Los siguientes datos representan el costo de energía eléctrica ($) durante el mes de Diciembre de 2017 de 50 casas que son alquiladas como oficinas en Posadas.

|  |  |
| --- | --- |
| **Costo de energía ($)** | **fa** |
| 800 – 990 | 4 |
| 990 – 1180 | 6 |
| 1180 – 1370 | 9 |
| 1370 – 1560 | 12 |
| 1560 – 1750 | 10 |
| 1750 – 1940 | 5 |
| 1940 – 2130 | 4 |
| Total | 50 |

1. Obtenga las medidas de tendencia central e interprételas.
2. Calcule los cuartiles e interprételos.
3. Construya un gráfico de caja. ¿Cómo se distribuyen los costos de las 50 casas?
4. Obtenga e interprete los percentiles 20, 40, 60 y 90 e interprételos.
5. Calcule el desvío estándar, coeficiente de variación y el coeficiente de asimetría, e interprételos.

Ejercicio Nº8: Analizando los siguientes diagramas que representan las edades de los empleados de dos empresas IT (Information Technology), ¿Qué información puede sacar de los diagramas de cajas?



Ejercicio 9: Dos investigadores están interesados en estudiar los hábitos de bostezo de los estudiantes en las clases. Ambos registran el tiempo (en minutos) que demoran en realizar el primer bostezo cada alumno desde que empieza la clase de los profesores. El gráfico muestra los tiempos que demoran en bostezar los alumnos con el profesor Acosta.

1. Gráfico, Gráfico de cajas y bigotes

   Descripción generada automáticamente¿Cuántos minutos tardan en dormirse la mitad de los alumnos?
2. ¿Qué porcentaje de alumnos se queda dormido después de los 17 minutos?
3. Los tiempos con el Profesor Britez son los siguientes:

10,5 11,3 11,9 12 12,3 12,3 12,5 12,7 13,4 13,7 13,8 14,2 14,8 15,1 15,3 16,7 16,8 18,8 20,8

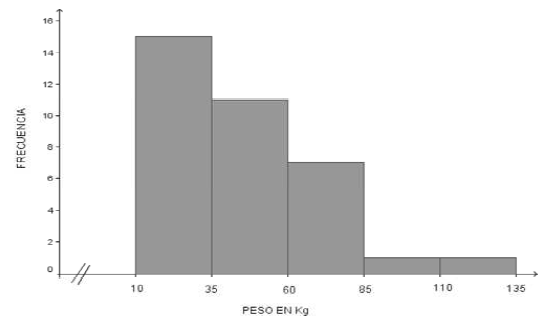
1. Construya un diagrama de caja correspondiente a los tiempos en la clase del Profesor Britez.
2. Realice un análisis comparativo de los diagramas.

Ejercicio Nº10: En un supermercado se realizó un estudio de eficiencia cuyos resultados representan los tiempos, en minutos, requerido para atender a 50 clientes en la caja:



1. ¿Cuál es el tiempo promedio de atención?
2. Describa el tiempo de atención usando la mediana.
3. ¿Cuál es el valor de D6?
4. ¿Cuál es el valor de Q3?
5. Interpreta las medidas obtenidas en los ítems anteriores.
6. El 30% de los clientes que menos esperó ser atendido en la caja, ¿cuánto tiempo esperó?, ¿y el 20 % de los que más esperó?
7. ¿Cuál medida, media o mediana, considera usted que es más representativa del conjunto de datos? Explique.
8. Realice un diagrama de caja y analice el mismo la forma de la distribución.
9. Determine el desvío estándar y el coeficiente de variación.
10. ¿Cuántos de los datos tienen valores entre  ¿qué porcentaje de la muestra representan?. Compara la respuesta obtenida con los resultados predichos por el teorema de Chebyshev.
11. ¿Es una distribución simétrica? en caso que no lo fuera explica que tipo de asimetría existe.
12. ¿cómo es el apuntamiento (kurtosis) de la distribución?

Ejercicio Nº11: Un grupo de especialistas en nutrición, médicos y preparadores físicos, idearon un plan para que personas obesas puedan bajar de peso de manera “natural”, es decir, sin suministrarles ningún tipo de droga ni sometiéndolos a excesivas rutinas de ejercicio. Para ello seleccionaron una muestra aleatoria de personas, con aproximadamente el mismo peso, que concurren a una clínica especializada en nutrición y le aplicaron el tratamiento durante seis meses. El siguiente gráfico indica los kilogramos perdidos por el grupo de pacientes con sobrepeso sometidos a la dieta especial



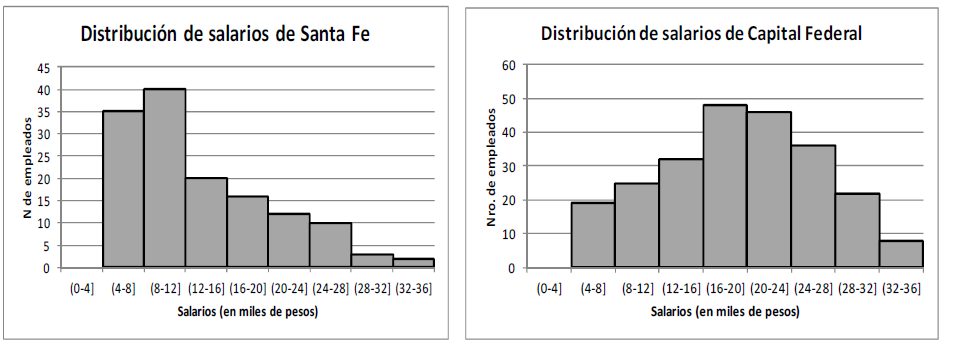
1. Un alumno decidió calcular la media aritmética y la calculó de la siguiente forma: ¿Es correcto lo realizado?
2. Calcula la medida de tendencia central y de dispersión más adecuada a esta situación. Justifica tu elección. Escribe una breve interpretación de ambas.
3. Describe las características principales (medidas de resumen) de esta distribución.

Ejercicio Nº12: En un ensayo de sanidad vegetal se estudia el efecto de una nueva formulación química para el control de una plaga en ajos blancos. Los datos se resumen en los gráficos que se presentan.

|  |  |
| --- | --- |
| a. Indica el tamaño de las muestras y la variable bajo estudio.  b. ¿En qué rango de valores se concentran los datos para ambos casos?  c. Indica qué medida de tendencia central elegirías en cada caso. Fundamenta tu elección.  d. Las distribuciones observadas, ¿tienen la misma variabilidad? Fundamenta tu respuesta.  e. ¿Qué podrías decir respecto a la asimetría de las gráficas? Corrobore analíticamente.  f. ¿Podrías hacer alguna conjetura sobre el efecto de la nueva formulación química? |  |

Ejercicio Nº13: En los histogramas que siguen, se presentan las distribuciones de salarios de los empleados de dos sucursales de una multinacional, una situada en la ciudad de Santa Fe y la otra, en Capital Federal en el año 2013. En base a la información presentada en estos gráficos, responde a las siguientes consignas:

Tamaños de muestras: 138 empleados de santa Fe y 234 empleados de Capital Federal



a) Analiza la simetría o asimetría de cada una de las distribuciones. Describe en qué información te basas para sacar tu conclusión.

b) Analiza las siguientes afirmaciones e indica si son o no correctas. En las que sean correctas, muestra la evidencia numérica que apoya tu afirmación. En las que sean incorrectas, reescríbelas de tal manera que sean correctas.

b.1. Si consideramos el 25% de los empleados con salarios más elevados en cada sucursal, podemos decir que el salario mínimo de ese grupo de empleados es mayor en Santa Fe que en Capital Federal.

b.2. Considerando nuevamente el 25% de los empleados con salarios más elevados, podemos indicar que en la sucursal de Santa Fe el número de empleados con esa característica es mayor que en la sucursal de Capital Federal.

b.3. Si se toma en cuenta el 50% central de los salarios, podemos indicar que en la sucursal de Capital Federal hay aproximadamente 117 empleados que cobran salarios entre 9 y 24 mil pesos aproximadamente, mientras que en la sucursal de Santa Fe, hay 70 empleados que ganan entre 10 y 17 mil pesos. En conclusión, si se pudiera elegir en qué sucursal trabajar, convendría hacerlo en la sucursal de Capital Federal.

b.4. El percentil 10 de la distribución de Santa Fe corresponde a 6.000 pesos aproximadamente y el percentil 10 de Buenos Aires corresponde a 7.000 pesos aproximadamente. Esto implica que, en la sucursal de Santa Fe, hay un 10% de empleados que ganan como máximo 6.000 pesos aproximadamente, mientras que en la sucursal de Capital Federal hay un 10 % de empleados que ganan como máximo 7.000 pesos aproximadamente.

b.5. El percentil 90 para Santa Fe corresponde a 19.000 pesos aproximadamente y el percentil 90 para Buenos Aires corresponde a 23.000 pesos aproximadamente. Esto significa, que en Santa Fe, el sueldo mínimo del 10% que gana más es 19.000 pesos y, que en Capital Federal, el sueldo máximo del 90% de empleados con menores sueldos es de 23.000 pesos.