



Tecnológico de Monterrey

Campus Estado de México

Actividad Evaluable 3: Análisis con vectores

Profesores

Jorge Adolfo Ramírez Uresti

Herramientas computacionales: el arte de la analítica

Arturo Barrios Mendoza		A01168331
Lucio Arturo Reyes Castillo		A01378985
Leyberth Jaaziel Castillo Guerra		A01749850

Fecha de Entrega:
7 de Mayo de 2023

Comenzamos el análisis buscando el número de datos ingresados, es decir la longitud del vector. Con esto sabemos que se representan los seguidores de 13 personas.

```
> length(seguidores)
[1] 13
```

Al hacer una suma del vector, podemos observar que el número total de seguidores de las 13 cuentas es de 4346.

```
> sum(seguidores)
[1] 4346
```

También podemos observar que el número más bajo de seguidores es de 72, mientras que el más alto es de 651.

```
> min(seguidores)
[1] 72
> max(seguidores)
[1] 651
```

La media es de 334 seguidores y la mediana es de 325.

```
> mean(seguidores)
[1] 334.3077
> median(seguidores)
[1] 325
```

Por otra parte, la varianza es muy grande, pues los datos no son muy cercanos entre sí.

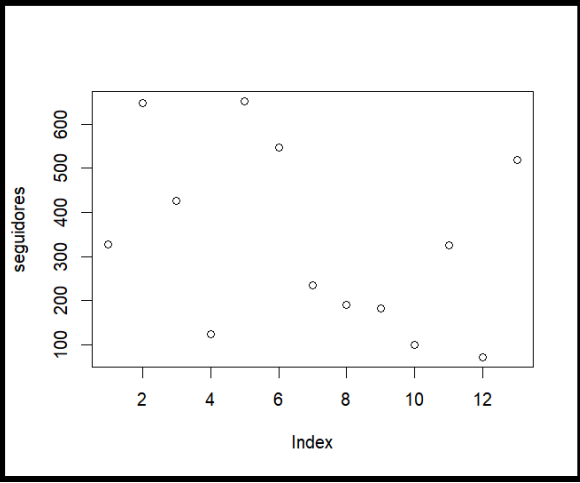
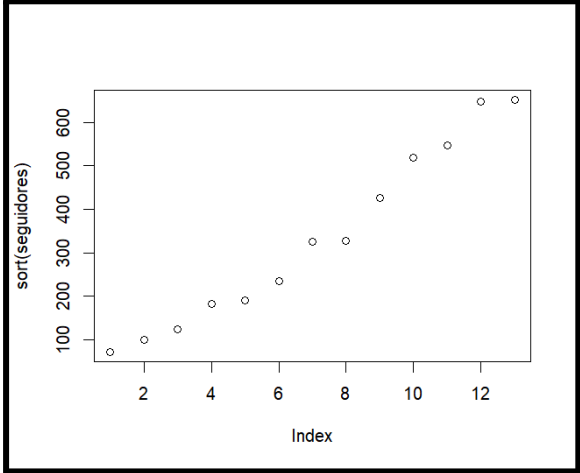
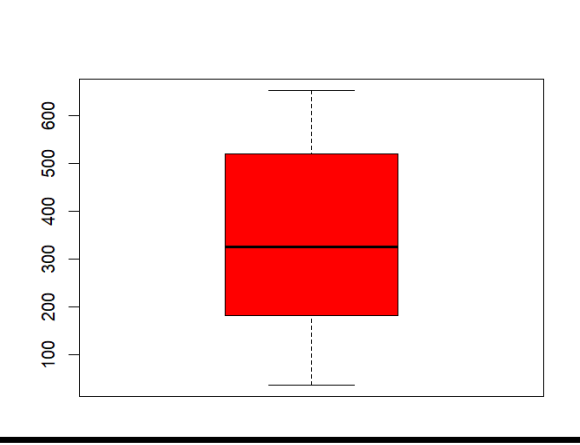
```
> var(seguidores)
[1] 42367.73
> sd(seguidores)
[1] 205.8342
```

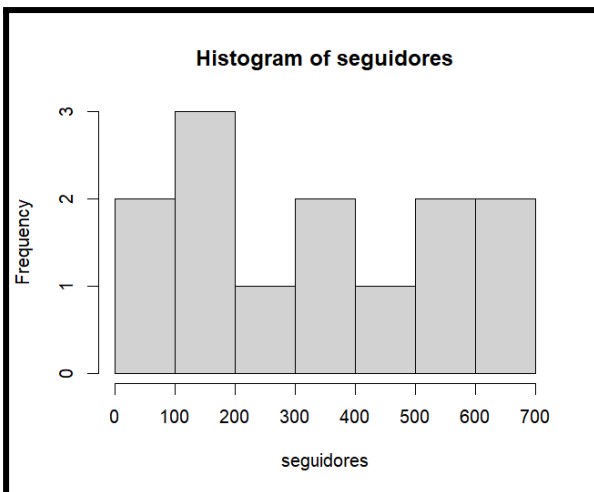
Al solicitar los datos únicos, nos damos cuenta de que no hay ningún valor que se repita; todos son distintos.

```
> unique(seguidores)
[1] 327 648 426 124 651 547 235 190 182 100 325 72 519
```

Con las siguientes líneas también podemos saber que el quinto valor es el más alto, mientras que el penúltimo valor es el más bajo.

```
> which.min(seguidores)
[1] 12
> which.max(seguidores)
[1] 5
```

Gráfica	Análisis
 <p>A scatter plot with 'Index' on the x-axis (values 2, 4, 6, 8, 10, 12) and 'seguidores' on the y-axis (values 100, 200, 300, 400, 500, 600). The data points are scattered, indicating they are not in order.</p>	<p>Se puede observar en la gráfica que los datos se encuentran en desorden y se muestran de la manera en que se obtuvieron de la base de datos. De tal manera que se deberá de realizar algún tipo de ordenamiento para lograr analizar los datos.</p>
 <p>A scatter plot with 'Index' on the x-axis (values 2, 4, 6, 8, 10, 12) and 'sort(seguidores)' on the y-axis (values 100, 200, 300, 400, 500, 600). The data points are sorted in ascending order, showing a clear upward trend.</p>	<p>Para observar la misma gráfica en orden se aplicó un sort() al vector original y posteriormente se graficó:</p> <pre data-bbox="855 1003 1378 1115">> sort(sort(seguidores)) [1] 72 100 124 182 190 235 325 327 426 519 547 648 651</pre>
 <p>A box plot with the y-axis ranging from 100 to 600. The box is red, with a black horizontal line representing the median at approximately 325. The whiskers extend from the box to the minimum and maximum values of the data.</p>	<p>A través de la gráfica de caja también podemos observar que la distancia entre los cuartiles es muy similar, lo que indica que los datos están dispersos, como ya había mostrado la varianza. Es posible observar también, que el tercer cuartil es más amplio. También se observa que no hay valores fuera de rango.</p>



En la gráfica podemos observar la cantidad de frecuencias que tenían los valores a través de un rango de 100 a 100 seguidores; se realizó de esta manera ya que los datos se repiten únicamente y no tendría tanto impacto. Sin embargo con esta forma de observar podemos hacer un análisis del mismo.

```
> summary(seguidores)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  72.0   182.0   325.0   334.3   519.0   651.0
```

Las gráficas muestran una diversidad en el número de seguidores, la variabilidad en este número sugiere que entre los compañeros existe una amplia gama de popularidades, lo que podría reflejar diferencias importantes en la actividad de cada persona en la plataforma, por ejemplo en la interacción con otros usuarios y el contenido compartido, un ejemplo es que la persona con 72 seguidores puede tener mucha menos actividad que la persona que tiene 651 seguidores. Otro aspecto que puede reflejar una visión similar es la dispersión de datos, la alta varianza y desviación estándar indica una amplia dispersión de los datos alrededor de la media, esto nos dice, de igual manera que el anterior punto, que mientras algunas personas tienen un número relativamente bajo de seguidores, otras tienen una cantidad considerablemente mayor lo que puede estar influenciado por factores como la actividad en la plataforma, popularidad personal y el tipo de contenido que se comparte. En conclusión, es posible realizar un análisis cualitativo a partir de datos cuantitativos, que nos revela que nuestros compañeros de clase muestran una diferencia significativa en su presencia o popularidad en instagram, puede proporcionar datos de la popularidad del uso de la plataforma en un grupo de compañeros, o en toda la escuela.

Sería interesante, por ejemplo, analizar la correlación entre el número de seguidores y otras variables como las calificaciones, la edad, o el número de horas que pasan en redes sociales.