

Lenguajes de programación - T08: Implementación de la distancia Damerau-Levenshtein en el lenguaje R

Jorge Aurelio Morales Manrique
C.C. 1010075711
jomorales@unal.edu.co

Universidad Nacional de Colombia
Abril 13 de 2021

7. Experimentación

A continuación se realiza la experimentación para 3 lenguajes de programación distintos: c++, java y python. A continuación se muestran los resultados de la ejecución del programa para cada caso.

C++

- Distancia promedio y varianza.

```
> main()
Read 63 items
[1] "Distancia promedio: " "6.30773169482847"
[1] "Varianza"           "4.72441252633611"
```

Figura 1: Media y varianza.

- Histograma.

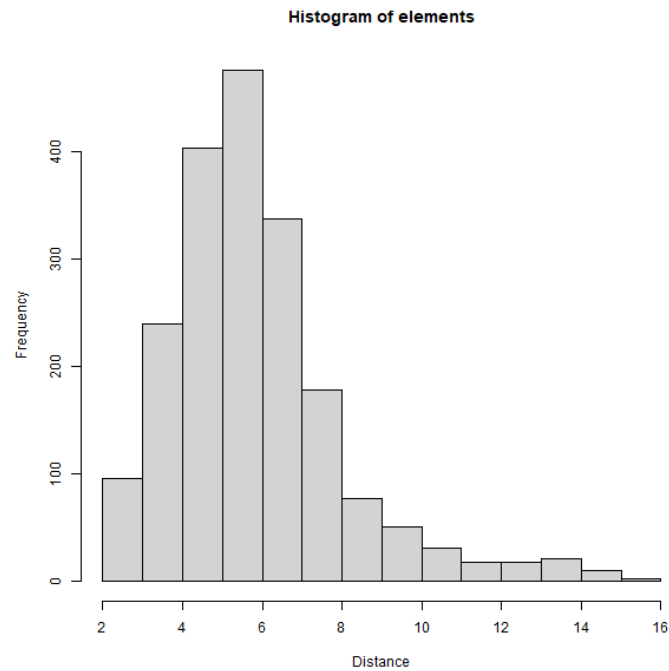


Figura 2: Histograma de las distancias.

Java

- Distancia promedio y varianza.

```
> main()
Read 50 items
[1] "Distancia promedio: " "6.09551020408163"
[1] "Varianza"           "3.1305762304922"
```

Figura 3: Media y varianza.

- Histograma.

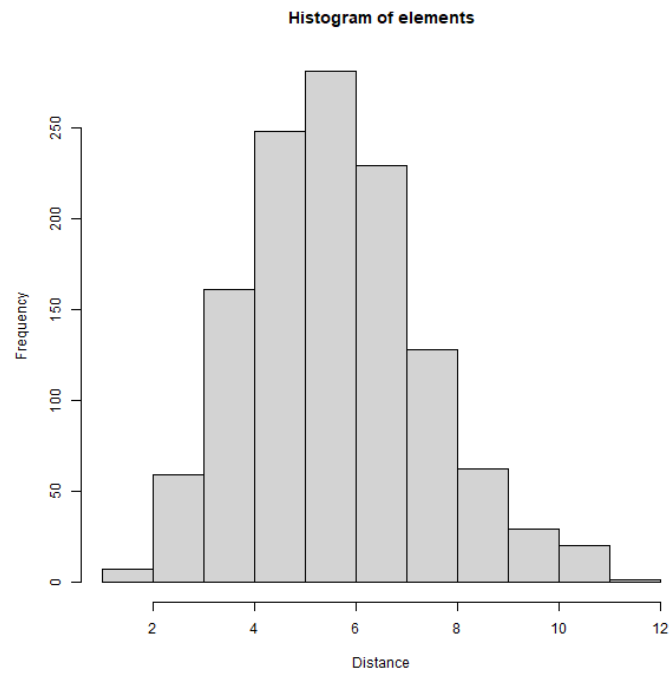


Figura 4: Histograma de las distancias.

Python

- Distancia promedio y varianza.

```
> main()
Read 35 items
[1] "Distancia promedio: " "4.81512605042017"
[1] "Varianza"            "2.0263701440172"
```

Figura 5: Media y varianza.

- Histograma.

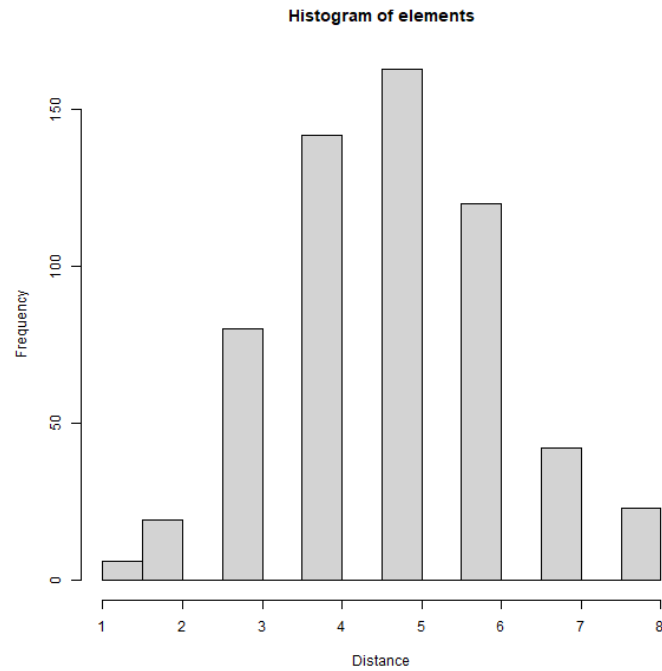


Figura 6: Histograma de las distancias.

8. Análisis de resultados

Se observa que C++ en comparación con Java y Python presenta la distancia promedio más grande, así como la varianza. En los 3 histogramas se obtiene una forma similar a la campana de Gauss, lo cual indica que los datos se podrían modelar utilizando una distribución normal. Al observar los histogramas generados se tiene que para C++ el intervalo de distancias es $[2, 15]$, para java $[1, 12]$ y para python $[1, 8]$.