**Hoja de Actividades**

Matemática y lingüística: La ley de Zipf

| La siguiente tabla muestra las diez primeras palabras de la Base de datos de la RAE:   | Base de datos de la Real Academia Española | | | | --- | --- | --- | | Ranking () | Palabra | Frecuencia () | | 1 | de |  | | 2 | la |  | | 3 | que |  | | 4 | el |  | | 5 | en |  | | 6 | y |  | | 7 | a |  | | 8 | los |  | | 9 | se |  | | 10 | del |  |   **¿Será posible estimar la frecuencia de una palabra dada su posición en el ranking?** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |

**Actividad 1**

1. Grafica la frecuencia de las primeras palabras y luego responde la siguiente pregunta: ¿Qué se puede observar respecto a la forma en que disminuye la frecuencia de las palabras a medida que se avanza en el ranking?
2. Responde las siguientes preguntas:
3. ¿Cómo varía la frecuencia a medida que se avanza en el ranking?
4. A partir del gráfico, ¿es posible apreciar la frecuencia de la palabra que está en la segunda posición del ranking? ¿Y en la posición ?
5. ¿Qué dificultad presenta este gráfico para visualizar los datos?
6. Completa la tabla con el logaritmo de la frecuencia, aproximado los valores a dos cifras decimales, y luego responde: ¿Cómo cambia la variable frecuencia al aplicar logaritmo?
7. ¿Qué tipo de función podría modelar el gráfico versus ?
8. Responde las siguientes preguntas:
9. ¿La pendiente del modelo es positiva o negativa? ¿Dónde cruza el eje ?
10. Observando los datos y la recta, ¿para qué rango de valores se ajusta mejor la recta?

**Actividad 2**

1. Completa la siguiente tabla:

| Ranking | Palabra | real | con el modelo | Diferencia positiva |
| --- | --- | --- | --- | --- |
|  | al |  |  |  |
|  | grupos |  |  |  |