



Análisis de tweets contexto peruano

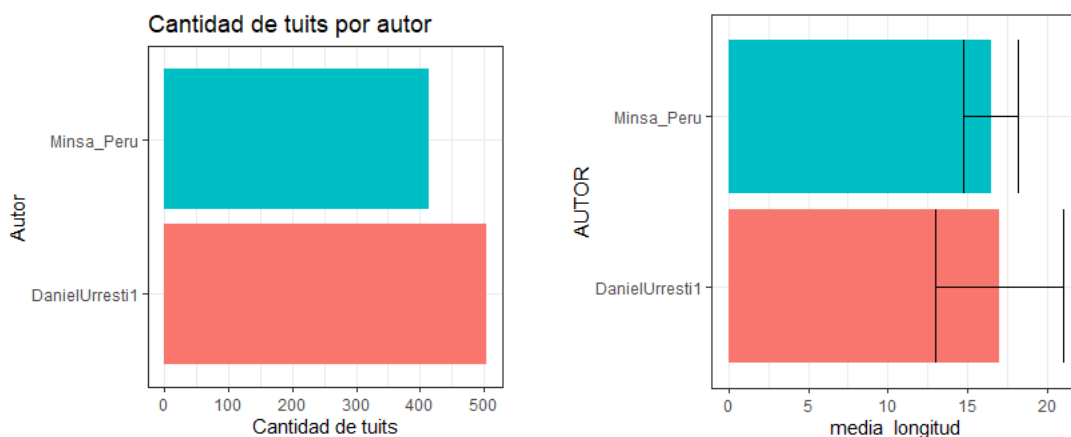
Alumnos: Leonela Tacure Purizaca

Se utilizó el conjunto de datos extraídos de twitter, se obtuvo 18 variables con 1,073 observaciones; al realizar el análisis se seleccionó solo 4 variables importantes:

- ✓ **ID:** Educación media de los titulares ocupacionales.
- ✓ **TEXTO:** Ingreso promedio en dólares.
- ✓ **FECHA:** Porcentaje de mujeres por ocupación.
- ✓ **AUTOR:** Prestigio de la ocupación, resultado de una encuesta social realizada a mediados

ANALISIS EXPLORATORIO:

1. Cantidad de tweets y frecuencia de palabras



AUTOR	n_tuits	palabras_utilizadas	palabras_diferentes	media_longitud	sd_longitud
DanielUrresti1	504	8.532	2.611	17,00	4,02
Minsa_Peru	414	6.804	1.644	16,40	1,73

En el dataset encontramos 504 tweets para DanielUrresti1 y 414 tweets para Minsa_Peru, al analizar las palabras utilizadas se encontró que Urresti utilizó mayor cantidad de palabras con 31% palabras diferentes en sus tweets mientras que Minsa_Peru tiene 24% de palabras diferentes. Respecto a Longitud media de los tweets por usuario, se observa que Minsa_Peru alterna más entre tweets cortos y largos, siendo su media inferior a los tweets de DanielUrresti1.

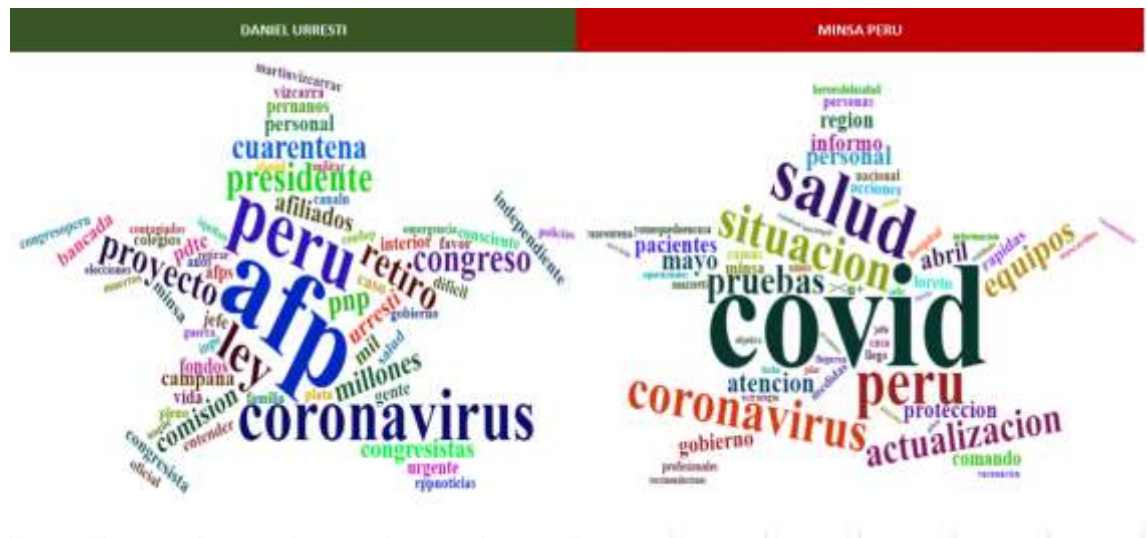
2. Palabras más utilizadas por Autor

Las palabras más utilizadas para Urresti son afp, peru, coronavirus, ley, retiro; se puede asumir por el contexto y el rango de fechas que se refiere al retiro de los fondos de las AFPs. En tanto para el MINSA son ministro, victor zamora, covid, salud, peru; por el contexto y las fechas mencionan con bastante énfasis al médico Víctor Zamora quien asumió el Ministerio de Salud los primeros meses de la pandemia.



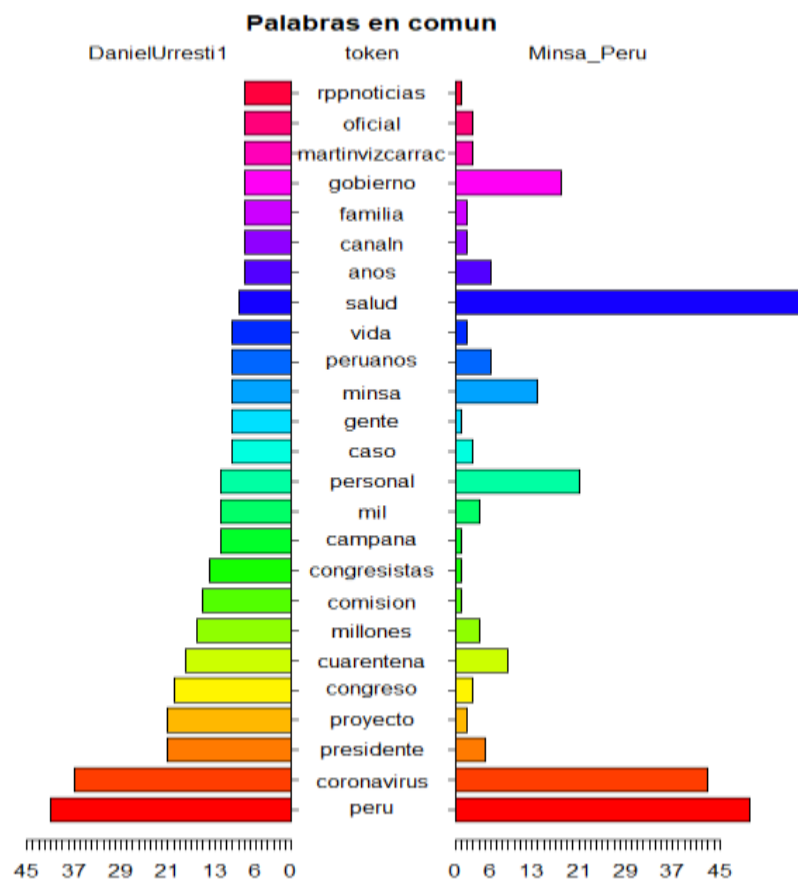
The image contains two horizontal bar charts. The top chart, titled 'DanielUrresti1', shows the top 5 terms for Daniel Urresti: 'afp' (approx. 65), 'peru' (approx. 42), 'coronavirus' (approx. 38), 'ley' (approx. 32), and 'retiro' (approx. 22). The bottom chart, titled 'Minsa_Peru', shows the top 5 terms for Minsa_Peru: 'covid' (approx. 65), 'salud' (approx. 58), 'peru' (approx. 45), 'situacion' (approx. 38), and 'coronavirus' (approx. 38). Both charts have a light gray background and white grid lines.

Term	Value (DanielUrresti1)	Value (Minsa_Peru)
afp	65	-
peru	42	45
coronavirus	38	38
ley	32	-
retiro	22	-
covid	-	65
salud	-	58
situacion	-	38



3. Palabras en común

Se encontró 302 palabras en común, destacando salud, peru, coronavirus, personal, gobierno, minsa. Se observa que las palabras que usa Urresti son utiliza las mismas palabras de modo más constante mientras el Minsa varia mucho en el lexico.





MODELAMIENTO: