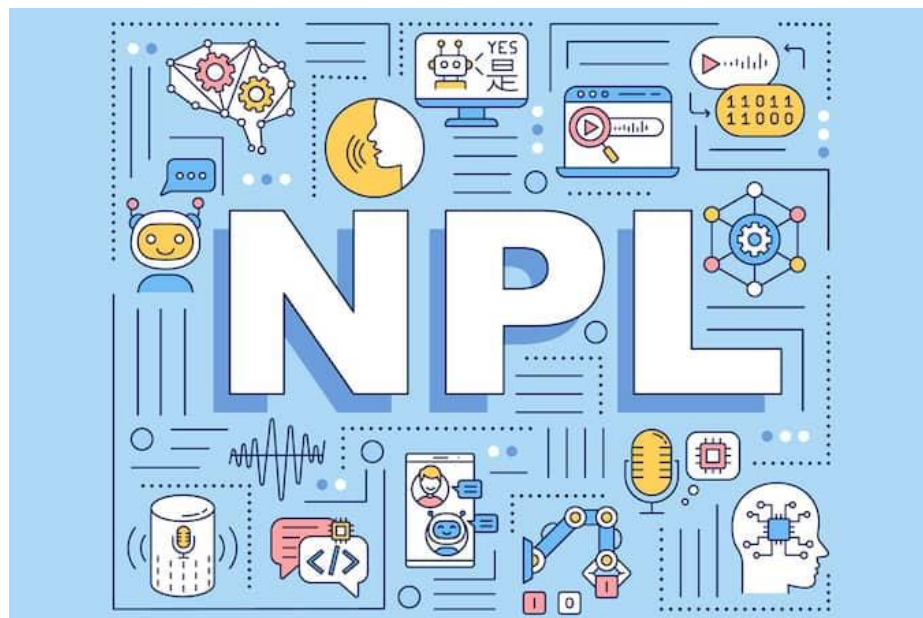




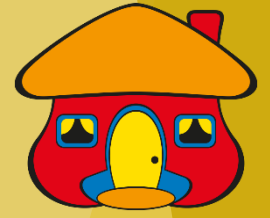
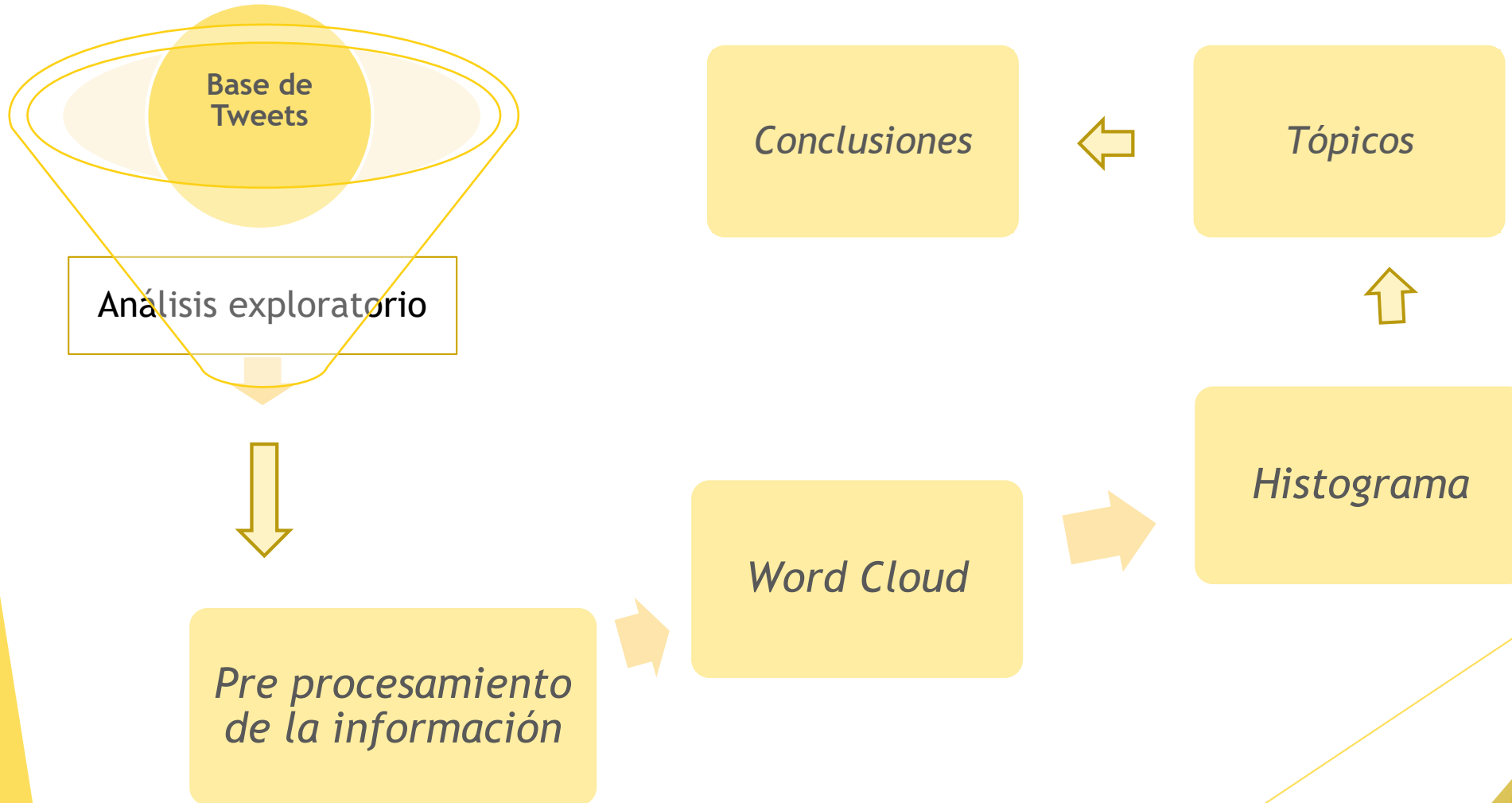
Análisis de Tweets relacionados con



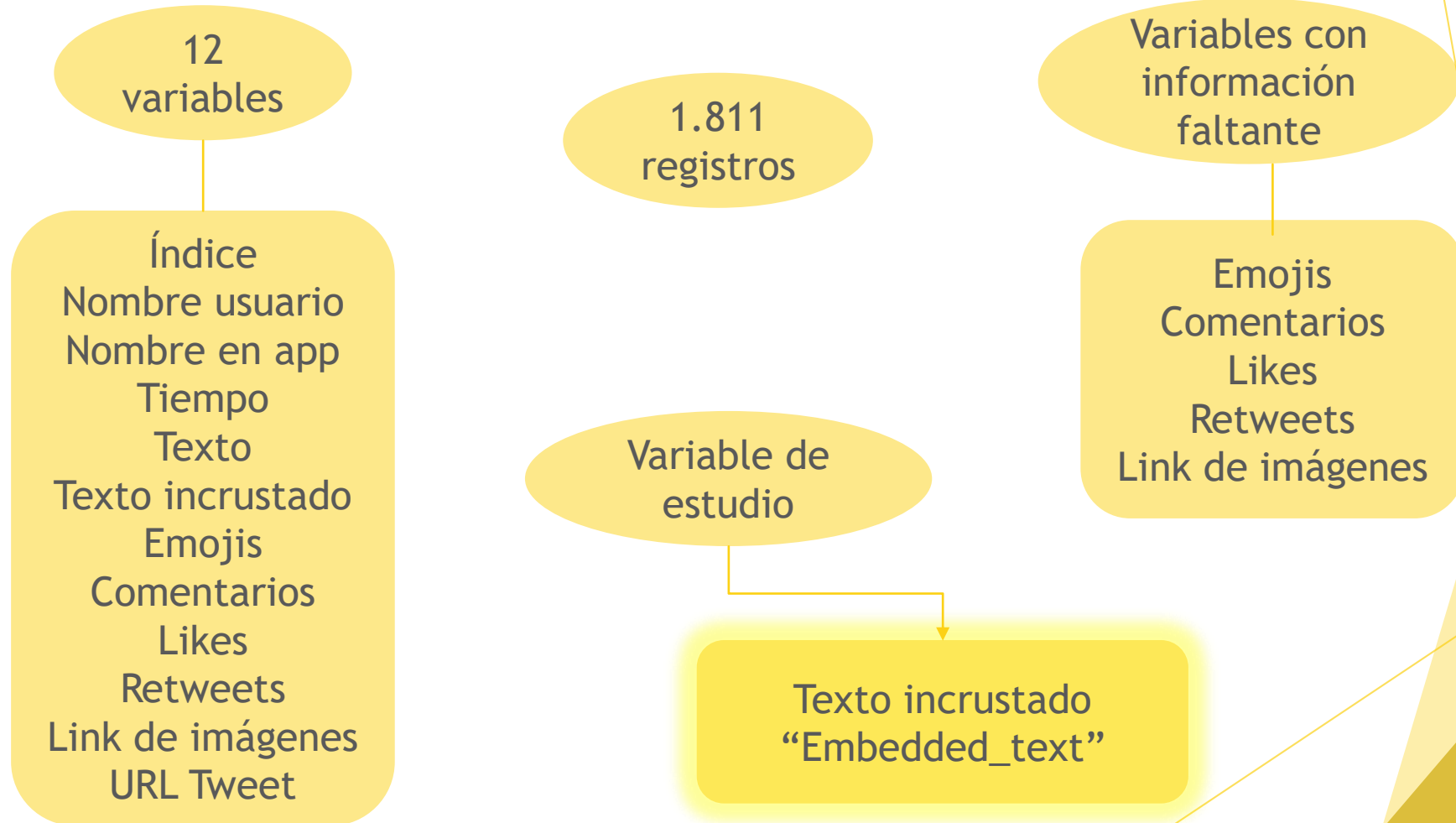
DAVIVIENDA



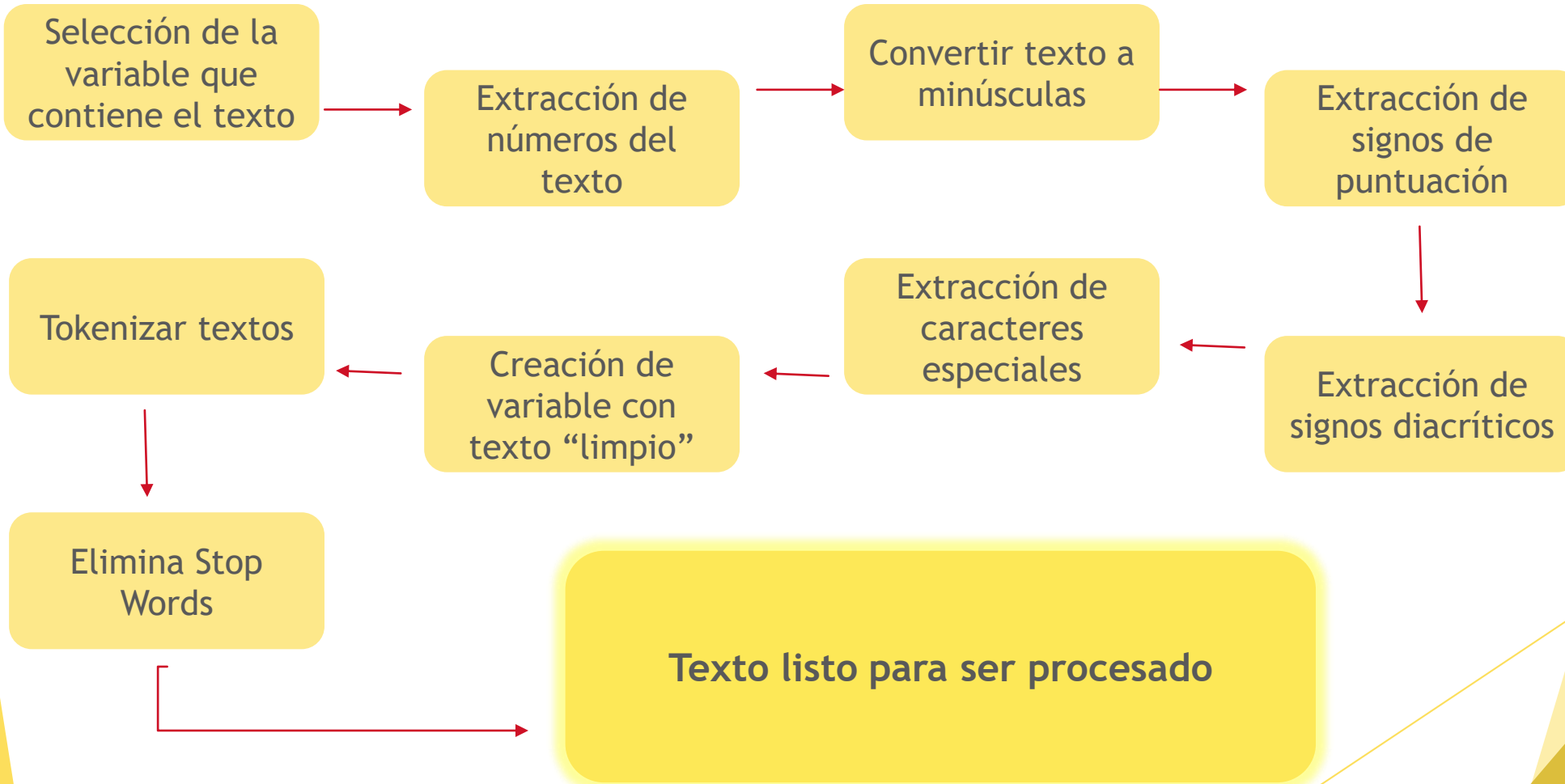
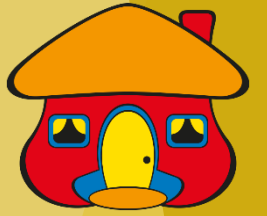
Desarrollo metodológico:



Análisis exploratorio de datos:



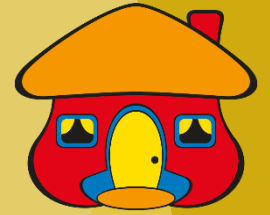
Pre procesamiento de la información:



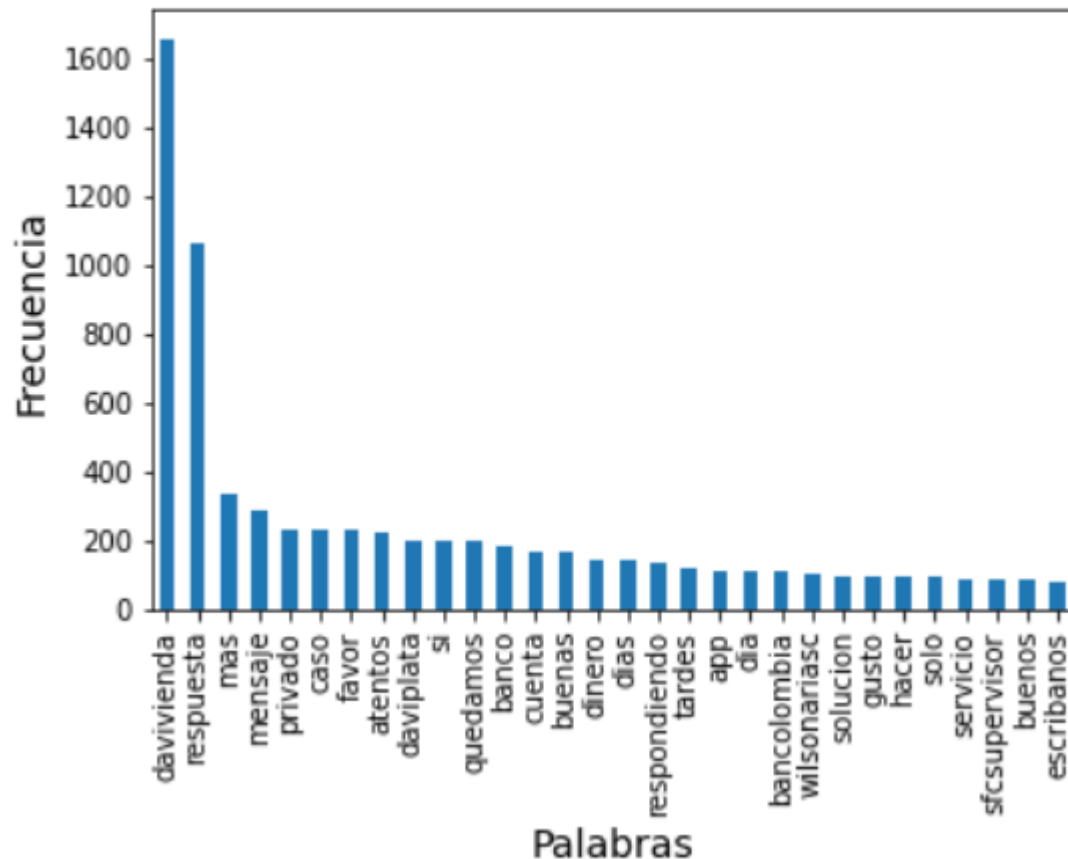
Word Cloud:



Histograma de palabras:



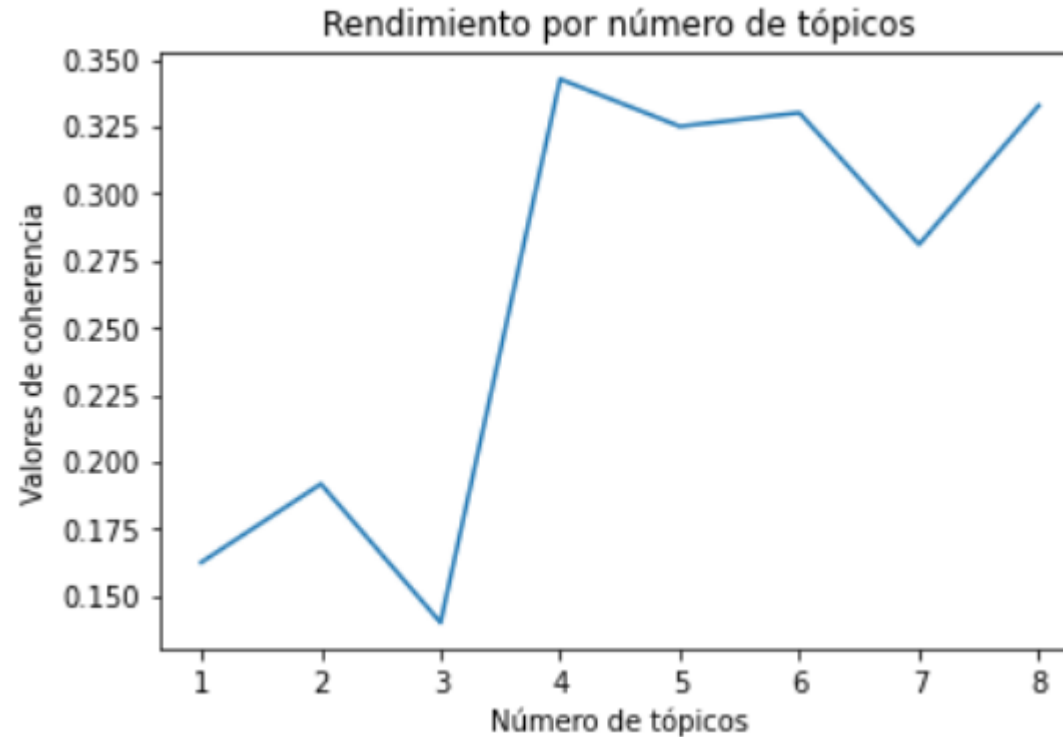
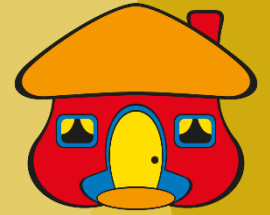
Histograma de las 30 palabras más relevantes encontradas en los Tweets



1. Davivienda.
2. Respuesta.
3. Mas.
4. Mensaje.
5. Privado.
6. Caso.
7. Favor.
8. Atentos.
9. Daviplata.
10. Si.

Se evidencia una serie requerimientos realizados por los usuarios del banco que esperan respuesta.

Número óptimo de tópicos:



Se usó un modelo de coherencia de la librería Gensim para obtener los valores de coherencia y de esta manera determinar que 4 tópicos logran describir el contexto general de los mensajes referentes a Davivienda.

Tópicos:

Tópico 1: Mensajes relacionados con usuarios que requieren respuesta sobre la aplicación Daviplata.

Daviplata

Tópico 2: Mensajes privados enviados a la entidad por parte de los usuarios.

Imbox

Tópico 3: Mensajes relacionados con situaciones presentadas en cuentas bancarias.

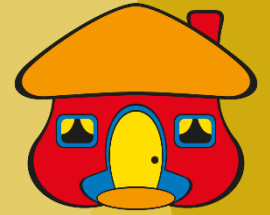
Cuentas
bancarias

Tópico 4: Mensajes relacionados con usuarios que solicitan respuestas sobre otros productos o servicios del banco.

Otros
productos



Conclusiones:



Los resultados presentados son de vital interés para el departamento de servicio al cliente, toda vez que evidencian una serie requerimientos realizados por los usuarios del banco que esperan pronta respuesta.

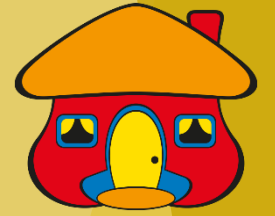
De acuerdo con el modelo no supervisado de tópicos se pudo determinar un total de 4 temáticas generales presentadas en los mensajes recolectados de Twitter; los cuales son app Daviplata, imbox a la entidad, cuentas bancarias y otros productos/servicios.

Las palabras más frecuentes del corpus fueron Davivienda, Respuesta, Mensaje, Privado, Quedamos, atentos; las cuales contextualizan un escenario de requerimientos hechos por usuarios de la entidad.

El análisis exploratorio permitió determinar que la data está conformada por 12 variables, 1.811 registros y que la variable objetivo era "Embedded_text".

Para obtener resultados óptimos en el procesamiento de lenguaje natural, es necesario realizar una adecuada limpieza a la información y preparar la data para su análisis.

Puesta en producción:



La interfaz creada para el usuario final se podría construir en Gradio y su desarrollo contendría los siguientes aspectos:

1. Guardar el modelo que realiza el Word cloud y los tópicos.
2. Instalar Gradio.
3. Definir una función que tenga como entrada el corpus objeto de estudio y retorne el Word Cloud, histograma de palabras más frecuentes y tópicos de acuerdo con el modelo construido.
4. Cargar la interfaz de Gradio con los respectivos parámetros.
5. Suministrar al usuario final la dirección URL donde puede acceder al modelo y hacer uso libre del mismo.



¡Gracias!