

Laboratorio 6.

Análisis de redes sociales

INSTRUCCIONES:

Debe extraer los datos de, al menos, una red social del tema de cada ejercicio. Debe hacer un análisis exploratorio para entender mejor los datos. Debe descubrir patrones interesantes. Recuerde que la investigación debe ser reproducible por lo que debe guardar el código que ha utilizado para resolver los ejercicios.

PROBLEMAS A RESOLVER

PROBLEMA 1.

Extraiga los datos relacionados con el tráfico en la ciudad de Guatemala, puede utilizar cualquier red social, a la que se tenga acceso, y pueden extraerse datos de varias. Por ejemplo, si extrae datos de twitter (ahora X) podría utilizar los hashtag #TraficoGT, #TransitoGT o @amilcarmontejo . Explore los datos que extraiga y descubra conocimiento, tendencias y elementos interesantes. ¿Cómo ha venido a complicar el tráfico en toda la ciudad la época de lluvia? ¿El socavón de zona 5 ha tenido un impacto importante en el tráfico de la zona de la universidad?

PROBLEMA 2.

Extraiga los datos que le permitan determinar la popularidad de Bernardo Arévalo y de Sandra Torres, puede utilizar cualquier red social, a la que se tenga acceso, y pueden extraerse datos de varias. Explore los datos que extraiga y descubra conocimiento, tendencias y elementos interesantes. ¿La opinión pública de cada uno de los candidatos es en su mayoría positiva o negativa? Puede usar comentarios para responderse esta pregunta. ¿Encontró algún influencer? ¿Cómo describiría la estructura de la red de cada uno de los candidatos, centralizada o distribuida? ¿Fue durante su campaña la red social un aspecto importante para influir en la decisión de los guatemaltecos?

EJERCICIOS

1. Descargue los archivos de datos
2. Cargue los archivos de datos a R o a Python.
3. Limpie y preprocese los datos. Describa de forma detallada las actividades de preprocesamiento que llevó a cabo.
 - 3.1. Se pueden hacer tareas como:
 - Convertir el texto a mayúsculas o a minúsculas
 - Quitar los caracteres especiales que aparecen como “#”, “@” o los apóstrofes.
 - Quitar las url
 - Revisar si hay emoticones y quitarlos, ¿conviene quitarlos para este ejercicio?
 - Quitar los signos de puntuación
 - Quitar los artículos, preposiciones y conjunciones (stopwords)

- Quitar números si considera que interferirán en los análisis.
- 4. Haga un análisis exploratorio de los datos para entenderlos mejor, documente todos los análisis. Escriba una serie de insights que se puedan seguir investigando.
- 5. Responda las preguntas que se le hacen en los ejercicios, las respuestas deben estar basadas en datos.
- 6. Llegue a conclusiones sobre lo que descubrió.

EVALUACIÓN

NOTA: La evaluación de cada integrante del grupo será de acuerdo con sus contribuciones al trabajo grupal

(15 puntos) Extracción de información:

- Está documentado el proceso de extracción de los datos de las redes sociales. Se extraen suficientes datos para descubrir información.

(25 puntos) Análisis exploratorio:

- Se elaboró un análisis exploratorio en el que se explican los cruces de variables, hay gráficos explicativos y análisis que permiten comprender el conjunto de datos.

(20 puntos) Limpieza y preprocesamiento de los datos:

- Se documentan las tareas de limpieza, incluyendo los paquetes/módulos que se usaron.

(30 puntos) Descubrimiento de información.

- Se describe de forma comprensible y amena los descubrimientos encontrados. Se calculan algunas métricas. Se detalla la información que es interesante. Se responden las preguntas formuladas de acuerdo con el análisis de los datos obtenidos de las redes sociales.

(10 puntos) Conclusiones.

- Se llega a conclusiones sobre los descubrimientos hechos utilizando los análisis.

MATERIAL A ENTREGAR

- Archivo .pdf con el informe que contenga, los resultados de los análisis y las explicaciones.
- Link de Google drive donde trabajó el grupo.
- Script de R (.r o .rmd) o de Python que utilizó.
- Link del repositorio usado para versionar el código.

FECHAS DE ENTREGA

- **AVANCE:** Prepare todo para descargar los datos de las redes sociales de ambos ejercicios, especifique perfiles, o cuentas que estará analizando. Descarga de los datos a trabajar. Tareas de limpieza sobre los datos. Planificación de las tareas para el grupo: viernes 8 de septiembre de 2023 a las 17:20.
 - **DOCUMENTO FINAL COMPLETO:** martes 12 de septiembre de 2023 a las 23:59
- NOTAS:**
- Solo se calificará el Documento Final si está entregado el avance con todo lo que se pide.
 - Solo los estudiantes que estén inscritos a algún grupo tendrán nota.