

- 1- José g
- 2- Cielo l

LA DELINCUENCIA Y SU RELACION CON LA ECONOMIA Y CORRUPCION

-1 Contexto y motivación: discuta la motivación y razones para su elección del proyecto, incluyendo el contexto y alguna investigación que haya incluido en su decisión. Comente también sobre la audiencia objetivo de su análisis (quién será el tomador de decisiones que se apoyará en su trabajo).

Respuesta: Las motivaciones principales por la cual elegimos abarcar la resolución del problema propuesto es porque, debido a la popular y real creencia de que las tasas de criminalidad, aquí en Chile, son bastante altas, y han estado en aumento, esto siendo evidenciado por las noticias matutinas que el grupo en general a visualizado, además de acontecimientos que demuestran la agravía de estos casos como: Quince homicidios se han contabilizado desde la noche del pasado viernes 12 de julio hasta la tarde de este martes 16 (1).

Ante esto nos preguntamos como equipo, ¿pueden estas ser ocasionadas a razón del “descontento” o no planeidad financiera general? ¿es la corrupción en las instituciones políticas un factor?

Esta información tendrá como audiencia objetivo todo aquel que esté interesado en saber cuáles pueden ser las causas hayan evidenciado una particular presencia de la delincuencia en ciertos lugares de la región metropolitana.

-2 Objetivos (2): describa los objetivos científicos y de inferencia de su proyecto. ¿Qué pregunta espera resolver o responder, y para qué? ¿Cuáles son los beneficios o accionables que podría lograr mediante este proyecto? ¿Cuál es la audiencia objetivo de su análisis?

Respuesta:

-3 Datos (1): describa los datos a utilizar, sus características (variables, tipo, formato, volumen), origen y forma de recolección. Los datos pueden provenir de fuentes privadas (siempre que no tengan restricciones de uso o publicación), de fuentes públicas (gobierno, plataformas de datos abiertos, web, etc.) o pueden ser recolectados mediante algún método propuesto por su equipo (Ej. web scrapping, Twitter, RRSS, etc.).

Respuesta: Los datos que se utilizarán para este trabajo serán principalmente de índole publica, provenientes de páginas web como el *CEAD*, *BANCO CENTRAL DE CHILE* e *INSTITUTO NACIONAL DE ESTADISTICA*, por mencionar algunos, para conseguir datos centralizados en el objetivo del proyecto. Gracias al ser en su mayoría paginas gubernamentales, una gran parte de los datos pueden ser descargar en formato Excel, o a su vez, utilizar métodos de ‘*WebScrapping*’ para

acceder a los datos. Cabe mencionar que este último no se podrá utilizar en todos los casos. Además de utilizar un repositorio de GitHub público (4) que tiene los datos geolocalizados de la región metropolitana de Chile, este caso es para mero temas ilustrativos. En resumen, se utilizará páginas gubernamentales y repositorios públicos, además de los métodos aprendidos en clase como 'WebScraping' para la realización del trabajo.

-4 Preguntas de investigación (2): proponga al menos 5 preguntas específicas que buscarán responder con base en los datos presentados. Estas preguntas pueden ser relativas a relaciones entre variables, capacidad de predicción de una variable, clasificación, identificación de clusters o categorías, etc.

Respuesta:

-5 Diseño tentativo (1): explique brevemente los métodos computacionales y estadísticos que espera usar en su análisis.

Respuesta: los métodos computacionales principales que creemos que utilizaremos para la realización de este proyecto serán los siguientes:

- **Librerías pandas (pd):** Con el objetivo de:

- 1- abrir `pd.read_[blank]()`
- 2- limpiar `df[df['variable'] == 'condicion']`
- 3- combinar `pd.merge() & pd.concat()`
- 4- entre otros `df.groupby() & df.fillna() & df.apply()`

Para una mejor comprensión y limpieza de los datos en los cuales trabajaremos.

- **Librerías matplotlib.pyplot (plt):** Con el objetivo de graficar los datos que tendremos y tener una comprensión más gráfica de los datos, con métodos como `plt.show()` o `plt.subplot()` por ejemplo.
- **Librerías request (r):** Con el objetivo de poder llevar a cabo el 'WebScraping' con métodos como `r.get()` y así aligerar el peso del repositorio y poner en práctica lo aprendido en clases.
- **Librerías pyarrow y pyarrow.parquet (pa, pq):** Con el objetivo de guardar los datos conseguidos y así no tener demoras o errores al tener que trabajar con datos ya cargados, con métodos como `pa.Table.From_pandas()` y `pq.write()`.
- **Librerías geopandas (gpd):** con el objetivo de leer y poder crear mapas de calor, por ejemplo, y así poder visualizar de mejor manera los datos regionales, con métodos como `gpd.read_file()`.

En resumen, se utilizarán la mayoría de los métodos aprendidos en clases para la realización de este proyecto.

Bibliografía:

- 1- <https://www.latercera.com/nacional/noticia/los-15-homicidios-ocurridos-en-los-ultimos-cuatro-dias-en-el-pais/JNPDIMCVV5ENNB5YNOQ6N5Y7QU/>
- 2- <https://tramitacion.senado.cl/appsenado/index.php?mo=transparencia&ac=doctoInformeAsesoria&id=25986>
- 3- https://obtienearchivo.bcn.cl/obtienearchivo?id=repositorio/10221/35843/2/BCN_GF_Evolucion_del_delito_en_Chile.pdf
- 4- <https://github.com/zorzalerrante/carto-en-python/blob/master/01%20-%20GeoPandas.ipynb>
- 5-