

Propuesta de Proyecto IMT 2200 2021-2

Grupo 8

Larry Uribe
Cristóbal Cancino
Benjamín Prieto
Diego Rodríguez

Github del proyecto: <https://github.com/08benja08p/Grupo-8-CDD>

Profesora: Dra. Paula Aguirre Aparicio.

Fecha: 15 de octubre de 2021.

Título del proyecto: Influencia de los factores sociales y económicos en la proyección de la delincuencia.

Contexto y motivación: El presente proyecto busca medir proyecciones sobre el aumento o disminución de los niveles de delincuencia por tipo de incidentes, considerando todas las comunas de Chile disponibles en la muestra, además se considera factores como índice de pobreza, población, superficie e incluso fase de paso a paso para los años de pandemia.

Objetivo general: Disponer de una mirada general país y también específica en delitos por comuna, en base a datos de fuentes oficiales para obtener una radiografía de la situación nacional que permita direccionar la toma de decisiones de forma comunal en materia social.

Audiencia del público: a la opinión pública (todo público) y organismos tanto públicos como privados para la toma de decisiones sociales relacionadas a la delincuencia.

Diseño tentativo:

1. Realizar un análisis descriptivo de los distintos datos a considerar.
2. Estudiar posibles correlaciones entre los datos.
3. Encontrar tendencias.
4. Graficar descriptivamente métricas relevantes.
5. El objetivo a largo plazo es poder entrenar un modelo que sea capaz de predecir el nivel de delincuencia en base a los datos considerados a lo largo del proyecto.

Preguntas Tentativas a Responder:

1. ¿Existe una relación entre los índices de delincuencia y la población en habitantes por metro cuadrado?
2. La delincuencia, ¿es más o menos violenta en relación a los niveles de pobreza por comuna?
3. ¿Cómo ha aumentado la delincuencia en todo el territorio nacional? ¿Se puede identificar grupos o clusters importantes?
4. ¿Son más comunas con mayor población las que poseen mayor índice de incidentes?
5. De acuerdo a los datos, ¿existe una relación entre delitos de mayor connotación social e índice de pobreza de la comuna?
6. ¿Existe alguna correlación simple entre la cantidad de delitos y alguna de las variables a analizar de los datos disponibles mencionados previamente?
7. En estos últimos años, en relación a las restricciones de movilidad impuestas por el plan paso a paso, ¿estas han afectado los índices y tipos de delitos en las distintas comunas y el país?

8. ¿Existen mayores delitos con resultado de muerte en comunas con menor acceso a sistemas de salud?
9. ¿El nivel de acceso a la educación influye en la cantidad de delitos de la comuna?

Datos:

- **Frecuencia Mensual_DENUNCIA.xlsx.** Corresponde a la frecuencia de delitos considerados de mayor connotación social, como son los robos, hurtos, lesiones, violaciones y homicidios. Esta data se encuentra subdividida por regiones y comunas. Obtenido del Centro de Estudios y Análisis del Delito (**CEAD**). Estos datos deben obtenerse manualmente de la aplicación del CEAD, ya que no cuenta con una API para la extracción de datos. Contiene datos desde el año 2005 al 2021, de frecuencia mensual.

```
RangeIndex: 421 entries, 3 to 423
Data columns (total 202 columns):
#   Column      Dtype
---  -
0   ORDEN       object
1   UN_ADMIN    object
2   UNIDAD      object
3   Período     object
4   Ene-2005    object
5   Feb-2005    object
6   Mar-2005    object
7   Abr-2005    object
8   May-2005    object
9   Jun-2005    object
10  Jul-2005    object
11  Ago-2005    object
12  Sep-2005    object
13  Oct-2005    object
```

Figura 1: Muestra de .info() de los datos para frecuencia delictiva.

Fuente: <http://cead.spd.gov.cl/centro-de-documentacion/>

Método de descarga: Manual en formato excel

- Datos sobre los nacimientos y las defunciones en Chile por comuna, para calcular la tasa de mortalidad y natalidad. Estos datos pueden ser encontrados en el GitHub del Ministerio de Ciencias. Se utilizarán datos para varios años con tal de ver la evolución de estas tasas.

```
RangeIndex: 17829 entries, 0 to 17828
Data columns (total 6 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Region          17829 non-null  object
1   Codigo region   17829 non-null  int64
2   Comuna          17829 non-null  object
3   Codigo comuna   17829 non-null  int64
4   Defunciones     17829 non-null  int64
5   Fecha           17829 non-null  object
dtypes: int64(3), object(3)
```

Figura 2: Muestra de .info() de los datos para defunciones 2020.

```

RangeIndex: 7893 entries, 0 to 7892
Data columns (total 6 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Region           7893 non-null   object
1   Codigo region    7893 non-null   int64
2   Comuna           7893 non-null   object
3   Codigo comuna    7893 non-null   int64
4   Nacimientos      7893 non-null   int64
5   Fecha            7893 non-null   object

```

Figura 3: Muestra de .info() de los datos para nacimientos 2021.

Fuente: <https://github.com/MinCiencia/Datos-COVID19/tree/master/input/RegistroCivil/>

Método de descarga: Directa a Jupyter desde repositorio github descrito previamente.

- **paso_a_paso.csv.** Información diaria del plan paso a paso por comuna, para estudiar la influencia de las cuarentenas en los niveles de delincuencia en el contexto de la pandemia.

```

RangeIndex: 386 entries, 0 to 385
Data columns (total 448 columns):
#   Column          Dtype
---  -
0   codigo_region    int64
1   region_residencia object
2   codigo_comuna    int64
3   comuna_residencia object
4   zona             object
5   2020-07-28       int64
6   2020-07-29       int64
7   2020-07-30       int64
8   2020-07-31       int64
9   2020-08-01       int64
10  2020-08-02       int64
11  2020-08-03       int64
12  2020-08-04       int64

```

Figura 4: Muestra de .info() de los datos para fases del plan paso a paso por comuna.

Fuente:

https://github.com/MinCiencia/Datos-COVID19/blob/master/input/Paso_a_paso/

Método de descarga: Directa a Jupyter desde repositorio el repositorio github descrito previamente.

- Directorio de establecimientos por año (csv). Datos sobre los establecimientos educacionales ingresados en el ministerio de educación. Obtenidos de la página de datos abiertos del mineduc. Se obtienen para varios años, con tal de obtener también la variación del número de establecimientos.

#	Column	Non-Null Count	Dtype
0	AGNO	16398 non-null	int64
1	RBD	16398 non-null	int64
2	DGV_RBD	16398 non-null	int64
3	NOM_RBD	16368 non-null	object
4	MRUN	16398 non-null	object
5	RUT_SOSTENEDOR	16398 non-null	object
6	P_JURIDICA	16398 non-null	int64
7	COD_REG_RBD	16398 non-null	int64
8	NOM_REG_RBD_A	16398 non-null	object
9	COD_PRO_RBD	16398 non-null	int64
10	COD_COM_RBD	16398 non-null	int64
11	NOM_COM_RBD	16398 non-null	object
12	COD_DEPROV_RBD	16398 non-null	int64
13	NOM_DEPROV_RBD	16398 non-null	object
14	COD_DEPE	16398 non-null	int64
15	COD_DEPE2	16398 non-null	int64
16	RURAL_RBD	16398 non-null	int64
17	LATITUD	16398 non-null	object
18	LONGITUD	16398 non-null	object
19	CONVENIO_PIE	16398 non-null	int64
20	ENS_01	16398 non-null	int64
21	ENS_02	16398 non-null	int64
22	ENS_03	16398 non-null	int64
23	ENS_04	16398 non-null	int64
24	ENS_05	16398 non-null	int64
25	ENS_06	16398 non-null	int64
26	ENS_07	16398 non-null	int64
27	ENS_08	16398 non-null	int64
28	ENS_09	16398 non-null	int64
29	ENS_10	16398 non-null	int64
30	ENS_11	16398 non-null	int64
31	MATRICULA	16398 non-null	int64
32	ESTADO_ESTAB	16398 non-null	int64
33	ORI_RELIGIOSA	16398 non-null	int64
34	ORI_OTRO_GLOSA	16398 non-null	object
35	PAGO_MATRICULA	16398 non-null	object
36	PAGO_MENSUAL	16398 non-null	object

Figura 5: Muestra de .info() de los datos para establecimientos educacionales al 2020.

Fuente:

<http://datos.mineduc.cl/dashboards/20015/descarga-bases-de-datos-directorio-de-establecimientos-educacionales/>

Método de descarga: Manual desde archivo exportado csv desde el portal previamente descrito.

- **InformacionComunas.csv.** Datos sobre la población comunal, obtenidos del GitHub del Ministerio de Ciencias.

```

RangeIndex: 346 entries, 0 to 345
Data columns (total 6 columns):
#   Column              Non-Null Count  Dtype
---  ---
0   Region              346 non-null    object
1   Codigo region       346 non-null    int64
2   Comuna              346 non-null    object
3   Codigo comuna       346 non-null    int64
4   Superficie_km2     346 non-null    float64
5   Poblacion           346 non-null    float64

```

Figura 6: Muestra de .info() de los datos para población al 2020.

Fuente:

<https://github.com/MinCiencia/Datos-COVID19/blob/master/input/Otros/InformacionComunas.csv>

Método de descarga: Directa a Jupyter desde el repositorio github descrito previamente.

- **Establecimientos.csv.** Datos sobre los establecimientos de salud reconocidos por el MINSAL al año 2020. Se utilizarán para cuantificar el nivel de acceso a salud en las diferentes comunas de Chile.

```

RangeIndex: 3814 entries, 0 to 3813
Data columns (total 7 columns):
#   Column                Non-Null Count  Dtype  
---  -
0   Región                3813 non-null   object  
1   Número de Región      3813 non-null   object  
2   Comuna                3813 non-null   object  
3   Establecimiento       3813 non-null   object  
4   Teléfono              2263 non-null   float64  
5   Pertenecencia         3813 non-null   object  
6   Dirección             2169 non-null   object  

```

Figura 7: Muestra de .info() para datos de hospitales por comuna al año 2020.

Fuente: <https://reportesdeis.minsal.cl/ListaEstablecimientoWebSite/>

Método de descarga: Manual desde la fuente especificada.