Logotipo

Descripción generada automáticamente

**UNIVERSIDAD PERUANA DE CIENCIAS APLICADAS**

FACULTAD DE INGENIERÍA

**SBA I\_FOUNDATIONS - USE OF BIG DATA**

**CASO DE ESTUDIO:**

**Relación entre el nivel educativo y los delitos denunciados en el Perú: Un análisis exploratorio de datos abiertos**

**GRUPO**

03

**SECCIÓN**

357

**ASESOR**

PALACIOS RUIZ, JULIO CESAR

**AUTOR(ES):**

* Ramos Peña, Aaron Daniel (U202419219)
* Rojas Fernandez, Juan Fernando (U202412532)
* Pajuelo Sifuentes, Percy (U20221A906)
* Salinas Martinez, Vicky Marion (U201823973)

**Lima, 23 de septiembre del 2025**

Contenido

[**I. Presentación de la empresa o tema para el proyecto:** 4](#_Toc209561729)

[**Motivos de investigación:** 5](#_Toc209561730)

[**II. Identificación del problema:** 5](#_Toc209561731)

[**Preguntas de investigación** 8](#_Toc209561732)

[**Pregunta principal:** 8](#_Toc209561733)

[**Preguntas secundarias:** 8](#_Toc209561734)

[**Objetivos** 8](#_Toc209561735)

[**Objetivo general:** 8](#_Toc209561736)

[**Objetivos específicos:** 8](#_Toc209561737)

[**III. Flujo de Información As Is:** 9](#_Toc209561738)

[**Diagrama Flujo de Información As Is** 10](#_Toc209561739)

[**IV. Arquitectura de datos As Is:** 10](#_Toc209561740)

[**Diagrama Arquitectura de datos As Is:** 11](#_Toc209561741)

[**Propuestas de mejora** 11](#_Toc209561742)

[**Fuentes:** 12](#_Toc209561743)

# **I. Presentación de la empresa o tema para el proyecto:**

El presente proyecto tiene como finalidad analizar la relación entre el nivel educativo de la población peruana y la incidencia de los delitos denunciados entre los años 2019 y 2023. El estudio se fundamenta en datos abiertos proporcionados por el Ministerio Público y el Instituto Nacional de Estadística e Informática (INEI), con el propósito de identificar patrones que permitan comprender si la educación constituye un factor asociado a la criminalidad. El tema reviste especial relevancia debido a que la inseguridad ciudadana se ha consolidado como uno de los principales problemas sociales en el país, mientras que la educación representa un determinante clave en el desarrollo humano y social.

**Definición de términos:**

* Nivel educativo: corresponde al máximo grado de instrucción alcanzado por la población de 15 años a más, clasificado en categorías tales como sin nivel/inicial, primaria, secundaria, superior no universitaria y superior universitaria (INEI, 2023).
* Delitos denunciados: comprenden las acciones u omisiones sancionadas por la ley penal, formalmente denunciadas ante el Ministerio Público. Incluyen, entre otros, violencia familiar, hurto agravado, abandono de familia, difamación y conducción en estado de ebriedad o drogadicción (Ministerio Público, 2023).

### **Motivos de investigación:**

La elección de este tema responde a la necesidad de comprender la interacción entre educación e inseguridad ciudadana. La educación, considerada un factor estructural del desarrollo, podría influir en la reducción de conductas delictivas; sin embargo, la relación no es unívoca y existen múltiples determinantes asociados al fenómeno criminal. A partir de este análisis, se busca proporcionar insumos que orienten políticas públicas en materia educativa y de seguridad.

# **II. Identificación del problema:**

El análisis exploratorio de datos muestra que, para el año 2023, el 46 % de la población peruana de 15 años o más había completado la secundaria, mientras que el 17.3 % alcanzó educación superior universitaria y el 15 % superior no universitaria. Sin embargo, existen marcadas diferencias regionales: en Cajamarca, solo el 37.4 % alcanzó la secundaria, mientras que en la provincia constitucional del Callao el 52.8 % completó dicho nivel.

En paralelo, los registros del Ministerio Público indican que los delitos más denunciados en el periodo 2019-2023 corresponden a violencia familiar, hurto agravado y omisión a la asistencia familiar. La persistencia de estos delitos refleja una problemática social de gran magnitud. Ante este contexto, surge la interrogante sobre si los niveles educativos inciden en la frecuencia y tipo de delitos cometidos.

A continuación, presentamos el nivel educativo del total de la población peruana mayor de 15:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **PERÚ: NIVEL DE EDUCACIÓN DE LA POBLACIÓN DE 15 Y MÁS AÑOS DE EDAD, SEGÚN DEPARTAMENTO, 2013-2023** | | | | | | | | | | | |
| (Porcentaje) | | | | | | | | | | | |
| **Nivel Educativo/ Departamento** | **2013** | **2014** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **2021** | **2022** | **2023** |
|  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Total** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** |  |
| Sin nivel/inicial | 4.4 | 4.3 | 4.1 | 4.2 | 4.0 | 3.8 | 3.7 | 3.4 | 3.5 | 3.5 | 3.3 |  |
| Primaria 1/ | 22.1 | 21.9 | 21.5 | 21.3 | 20.9 | 20.7 | 20.1 | 19.7 | 20.0 | 18.9 | 18.6 |  |
| Secundaria | 43.8 | 44.4 | 45.4 | 44.2 | 44.5 | 44.2 | 43.9 | 45.8 | 46.4 | 46.8 | 46.0 |  |
| Superior no universitaria | 13.5 | 12.9 | 12.5 | 13.6 | 13.5 | 13.8 | 14.4 | 14.3 | 14.1 | 14.5 | 15.0 |  |
| Superior universitaria 2/ | 16.3 | 16.5 | 16.4 | 16.7 | 17.1 | 17.6 | 17.8 | 16.7 | 16.0 | 16.4 | 17.3 |  |

Como podemos ver, para el 2023 el 46% de la población mayor a 15 había finalizado la secundaria. Estos datos fluctúan mucho dependiendo de la región, donde en Cajamarca podemos ver que el porcentaje baja hasta 37.4, mientras que en la provincia constitucional del Callao alcanza el 52.8%.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Cajamarca** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** |
| Sin nivel/inicial | 9.3 | 9.8 | 8.6 | 10.0 | 8.7 | 7.9 | 8.2 | 7.7 | 7.0 | 8.0 | 7.7 |
| Primaria 1/ | 40.7 | 40.7 | 40.2 | 41.3 | 41.6 | 41.2 | 37.9 | 37.2 | 36.4 | 35.3 | 33.9 |
| Secundaria | 32.0 | 32.5 | 33.9 | 33.0 | 34.2 | 31.9 | 35.1 | 35.3 | 35.4 | 36.3 | **37.4** |
| Superior no universitaria | 8.4 | 8.2 | 7.8 | 7.1 | 6.9 | 8.6 | 7.7 | 9.0 | 10.1 | 9.5 | 10.0 |
| Superior universitaria 2/ | 9.6 | 8.7 | 9.4 | 8.6 | 8.6 | 10.4 | 11.2 | 10.8 | 11.0 | 10.8 | 10.9 |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Prov. Const. del Callao** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** | **100.0** |
| Sin nivel/inicial | 1.3 | 1.3 | 1.8 | 1.7 | 1.6 | 1.4 | 1.5 | 1.4 | 1.7 | 1.8 | 1.1 |
| Primaria 1/ | 13.6 | 12.0 | 12.3 | 11.2 | 10.9 | 11.6 | 10.1 | 10.9 | 10.4 | 11.6 | 10.8 |
| Secundaria | 52.0 | 54.8 | 56.7 | 51.5 | 52.4 | 52.9 | 51.3 | 53.9 | 55.1 | 54.8 | **52.8** |
| Superior no universitaria | 18.6 | 16.7 | 15.7 | 20.3 | 18.4 | 19.2 | 19.5 | 18.2 | 17.4 | 17.0 | 18.8 |
| Superior universitaria 2/ | 14.6 | 15.2 | 13.5 | 15.3 | 16.7 | 14.9 | 17.7 | 15.7 | 15.4 | 14.8 | 16.4 |

Por otro lado, podemos ver que el porcentaje de personas mayores a 15 presenta una tendencia positiva en los últimos 11 años:

Este gráfico nos indica que hay un crecimiento promedio de 0.5% por año, reflejando una pequeña pero sostenida mejora.

Al analizar los delitos más comunes, conseguimos la siguiente información:



Podemos ver que el delito más común a nivel nacional en los últimos 5 años es la violencia familiar, seguido por el hurto agravado y luego varía entre omisión por pensión de alimentos o conducción por estado de ebriedad o drogadicción.  
  
A partir de esta información deseamos analizar si existe una correlación entre el nivel de educación y los delitos más comunes en este periodo de tiempo. Planteamos entonces las siguientes preguntas para la investigación:

## **Preguntas de investigación**

### **Pregunta principal:**

¿Existe una relación entre el nivel educativo de la población y la incidencia de los delitos más comunes denunciados en el Perú durante 2019-2023?

### **Preguntas secundarias:**

* ¿Qué delitos concentran la mayor cantidad de denuncias en el país y cómo han variado en el tiempo?
* ¿Cuál ha sido la evolución de los distintos grados de instrucción entre 2013 y 2023?
* ¿Es posible estimar la evolución futura de los delitos predominantes mediante análisis predictivo?
* ¿Qué tan determinante es el nivel educativo en la criminalidad y qué otros factores explican los delitos?

## **Objetivos**

### **Objetivo general:**

Analizar la relación entre el nivel educativo de la población peruana mayor de 15 años y los delitos denunciados a nivel nacional durante el periodo 2019-2023.

### **Objetivos específicos:**

* Identificar los delitos con mayor incidencia en el país en los últimos cinco años.
* Describir la evolución del nivel educativo de la población entre 2013 y 2023.
* Explorar posibles correlaciones entre nivel educativo y criminalidad.
* Proponer recomendaciones de política pública para reducir los delitos más recurrentes.

El objetivo del proyecto finalmente es analizar la relación entre el nivel educativo de la población mayor de 15 años y los delitos denunciados a nivel nacional, para finalmente presentar recomendaciones a las distintas regiones del Perú sobre políticas públicas que permitan combatir los delitos más recurrentes.

# **III. Flujo de Información As Is:**

Actualmente, el proceso de análisis se desarrolla mediante la recopilación de datos en fuentes oficiales. El Ministerio Público publica anualmente los delitos denunciados en formato CSV, mientras que el INEI dispone de registros históricos del nivel educativo en formato Excel. Los investigadores descargan manualmente los archivos, realizan la limpieza de los encabezados, consolidan la información y aplican herramientas básicas como tablas dinámicas para explorar la relación entre variables. Finalmente, los resultados se presentan en gráficos y reportes en Word o PowerPoint.

1. Generación de datos

* Ministerio Público → registra los delitos denunciados (CSV por año).
* INEI → recopila estadísticas de nivel educativo (Excel anual consolidado).

1. Publicación

* Ambas instituciones publican sus datos en plataformas de datos abiertos (web).

3. Extracción manual

* El equipo del proyecto descarga los archivos (CSV y Excel).

4. Preparación y limpieza

* Se unifican los CSV de delitos (2019-2023).
* Se corrigen nombres de encabezados.
* Se formatean datos en tablas dinámicas.

5. Análisis y visualización

* Se generan gráficos de tendencia y comparativos.
* Se cruzan variables de educación vs. delitos.

6. Reporte final

* Se presenta un informe con tablas, gráficos y un análisis predictivo básico.

## **Diagrama Flujo de Información As Is**

Gráfico, Diagrama

El contenido generado por IA puede ser incorrecto.

# **IV. Arquitectura de datos As Is:**

La arquitectura actual es manual y basada en archivos. Se compone de dos fuentes principales: (i) archivos CSV de delitos proporcionados por el Ministerio Público y (ii) un archivo Excel del INEI que consolida datos de educación. El procesamiento de datos se realiza en Excel, mediante limpieza, organización y uso de tablas dinámicas. El almacenamiento se limita a carpetas locales y el consumo se materializa en reportes estáticos. Esta arquitectura carece de automatización, escalabilidad y centralización, limitando la eficiencia del análisis.

* Fuentes de datos
  + CSV (delitos) → Ministerio Público (varios archivos).
  + Excel (educación) → INEI (un solo archivo con series históricas).
* Integración de datos
  + No existe sistema automatizado.
  + Se hace en Excel con tablas dinámicas.
* Almacenamiento
  + Los archivos se mantienen en carpetas locales.
  + No hay repositorio centralizado.
* Consumo
  + El análisis se hace en Excel y se presenta en PowerPoint/Word.
  + No hay dashboards en tiempo real ni herramientas de BI avanzadas.

## **Diagrama Arquitectura de datos As Is:**

Diagrama

El contenido generado por IA puede ser incorrecto.

### **Fuentes:**

Para obtener los delitos denunciados utilizamos los datos publicados por el Ministerio Público en <https://datosabiertos.gob.pe/dataset/mpfn-delitos> , desde el año 2019 al 2023. Los datos están en formato CSV, cada archivo es un año diferente.

Para los datos del nivel de educación utilizamos los datos publicados por el INEI en la página  
<https://m.inei.gob.pe/estadisticas/indice-tematico/education/> , donde tenemos los datos desde el 2013 hasta el 2023. Los datos están presentados en una sola hoja de Excel.

Para los datos del ministerio público es necesario agrupar todos los archivos CSV para poder trabajar con la data de todos los años. Luego es necesario limpiar los nombres de los encabezados de los datos y poner la información en una tabla dinámica.

La información finalmente se presenta como un reporte con gráficos y tablas que demuestren la relación entre la data y presenten también un análisis predictivo que facilite la toma de decisiones de las regiones a quienes se les presentará el reporte