



## **HERRAMIENTAS PARA INTELIGENCIA ARTIFICIAL**

Trabajo Final

**Ecuador / mayo 2024**

### **Integrantes:**

- **HENRY GUARNIZO**
- **EDUARDO ARAUJO**
- **ASTRID MEDINA**

Contenido

Generación de una aplicación de inteligencia artificial ..... 3

1. INTRODUCCIÓN ..... 3

2. MARCO TEÓRICO DE LAS TECNOLOGÍAS/LIBRERÍAS USADAS..... 3

3. DESCRIPCIÓN DEL DATASET USADO..... 3

4. DESCRIPCIÓN DE LOS PASOS REALIZADOS EN EL PROYECTO..... 3

    4.1 Descripción de las visualizaciones generadas ..... 4

5. CONCLUSIONES ..... 9

6. BIBLIOGRAFÍA ..... 10

7. ANEXOS ..... 10

# Generación de una aplicación de inteligencia artificial

## 1. INTRODUCCIÓN

Este proyecto analiza la distribución y tendencias de estudiantes en las instituciones educativas de Ecuador, usando datos históricos desde 2009 hasta 2024. Se investiga la evolución de la inscripción estudiantil para identificar patrones que podrían darnos un análisis para la toma de decisiones académicas y administrativas en el ámbito educativo.

## 2. MARCO TEÓRICO DE LAS TECNOLOGÍAS/LIBRERÍAS USADAS

Para este proyecto se han utilizado las siguientes librerías:

- **Pandas:** utilizada para la manipulación y el análisis de datos. Permite la manipulación de estructuras de datos con alto rendimiento y eficiencia.
- **SQLite:** sistema de gestión de bases de datos que facilita el almacenamiento y la consulta de datos en un formato relacional.
- **Matplotlib:** librería de visualización que permite crear una amplia gama de gráficos y visualizaciones estáticas, animadas e interactivas.
- **Bokeh:** para desarrollar visualizaciones interactivas, permitiendo una exploración dinámica de los datos.
- **Pygwalker:** procesamiento y análisis de datos.

## 3. DESCRIPCIÓN DEL DATASET USADO

En este trabajo se utilizó dos datasets que incluyen información académica y administrativa, tanto de inicio y fin de cada año lectivo. Contienen registros detallados de estudiantes, docentes, y características de instituciones educativas desde 2009 a 2024 en Ecuador. Estos datos incluyen:

- Datos demográficos y académicos por provincia y año.
- Información sobre el sostenimiento (público o privado), tipo de educación, y jurisdicción de cada institución.
- Estudiantes matriculados, promovidos, no promovidos y aquellos que abandonan sus estudios.

## 4. DESCRIPCIÓN DE LOS PASOS REALIZADOS EN EL PROYECTO

Para el proyecto se han realizado los siguientes pasos:

1. **Preprocesamiento de datos:** limpieza y preparación de los datos, incluyendo la detección de codificaciones, eliminación de filas mal formateadas, y división de columnas para una mejor manipulación.
2. **Creación de nuevas columnas:** se crearon cinco columnas para complementar el análisis que fueron:

- Modalidad: se agregó información sobre la modalidad educativa de cada institución.
- Jornada: información sobre el tipo de jornada escolar.
- Acceso\_Edificio: datos sobre la accesibilidad física de las instalaciones.
- Promovidos: cantidad de estudiantes promovidos cada año.
- Abandono: cantidad de estudiantes que abandonaron.

### 3. Fusión de datos

- Merge de datasets, se combinaron los datasets con formato .csv y .sql con base a las claves primarias Anio\_lectivo y AMIE (identificación de cada institución educativa).
- Se agruparon datos por provincias y años.

4. **Gráficas y visualizaciones:** Con el uso de las librerías como matplotlib, bokeh y pygwalker se realizaron gráficas que permitirían interpretar de mejor manera la información que contienen los datasets.

#### 4.1 Descripción de las visualizaciones generadas

**Gráfico de barras (Matplotlib):** muestra la cantidad total de estudiantes por provincia, facilitando la comparación regional. El eje X (Horizontal) representa las diferentes provincias. El eje Y (Vertical) indica la cantidad total de estudiantes en cada provincia, en millones. Por ejemplo, una barra que alcanza hasta  $1e7$  indica que esa provincia tiene 10 millones de estudiantes.

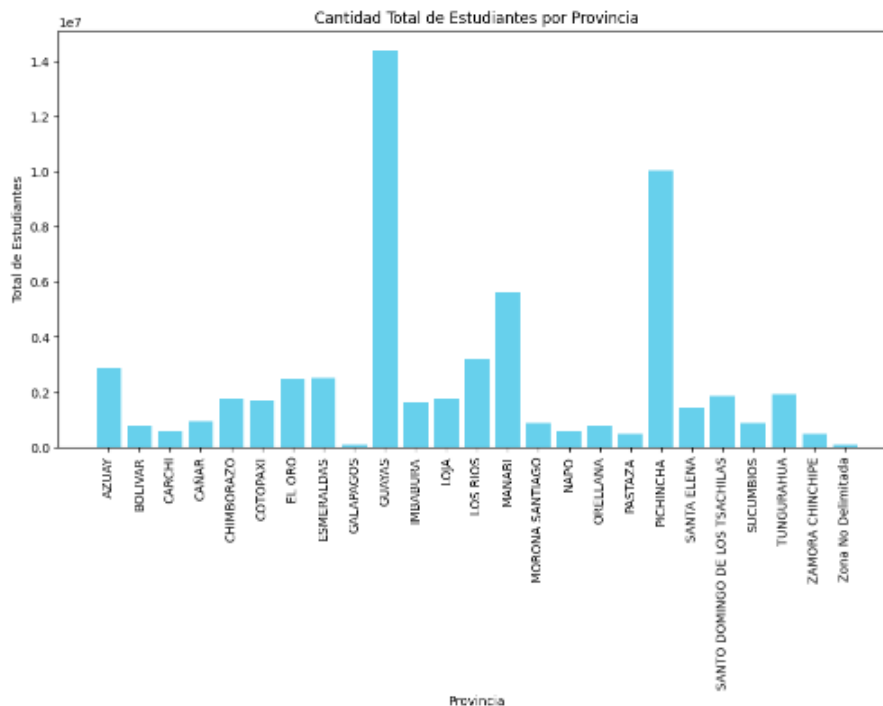
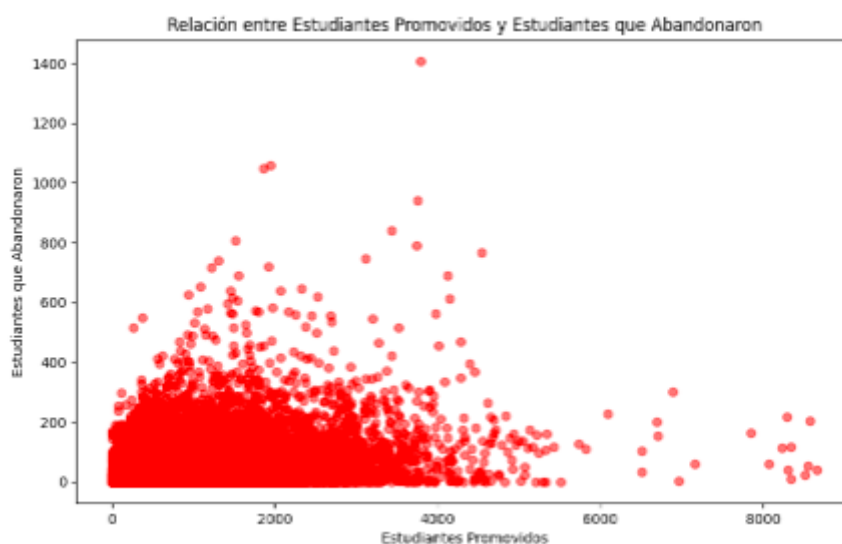


Figura 1: Total de estudiantes por provincia

### Hallazgos:

- En el gráfico se puede visualizar una distribución desigual de estudiantes entre las provincias. Algunas provincias, como Guayas y Pichincha, tienen una cantidad significativamente mayor de estudiantes en comparación con otras provincias como Bolívar o Cañar.
- Las provincias de Guayas y Pichincha resaltan con barras mucho más altas, indicando que estas provincias albergan una gran cantidad de estudiantes. Esto podría estar relacionado con la presencia de grandes ciudades y centros educativos en estas áreas.
- Las Provincias como Zamora Chinchipe y otras al final del gráfico muestran barras relativamente bajas, lo que indica una menor población estudiantil.

**Diagrama de dispersión (Matplotlib):** Se muestra la relación entre estudiantes promovidos y aquellos que abandonaron, destacando áreas críticas de atención. El eje X (Horizontal) representa la cantidad de estudiantes promovidos. La escala va de 0 a aproximadamente 8000 estudiantes. El Eje Y (Vertical) indica la cantidad de estudiantes que abandonaron, con valores que van desde 0 hasta más de 1400.



*Figura 2: Relación entre estudiantes promovidos y que abandonaron*

### Hallazgos:

- La mayoría de los puntos están concentrados cerca del origen, indicando que para muchos registros hay un número bajo tanto de estudiantes promovidos como de estudiantes que abandonaron.
- A medida que aumenta el número de estudiantes promovidos, se observa una dispersión más amplia en el número de estudiantes que abandonaron. Esto sugiere que en entornos donde hay más estudiantes promovidos, también puede haber más variabilidad en el número de abandonos.
- No parece haber una tendencia clara o lineal en el gráfico, lo que podría indicar que no hay una relación directa y simple entre el número de estudiantes promovidos y los que

abandonan. Sin embargo, hay una leve tendencia a que los abandonos aumenten con el aumento de promovidos en algunos casos.

**Gráficos interactivos (Bokeh):** Este tipo de gráfico nos permitió visualizar la distribución de una variable numérica a lo largo de varias categorías, en este caso, años lectivos, y cómo se descompone esa variable en subcategorías, aquí representadas por diferentes colores para cada año. El eje X (Horizontal) muestra diferentes provincias. Cada barra en el gráfico representa una provincia específica de Ecuador. El eje Y (Vertical) representa la cantidad total de estudiantes, escalada en notación científica. Cada color dentro de una barra representa un año lectivo diferente, comenzando en 2009-2010 y continuando hasta 2023-2024.

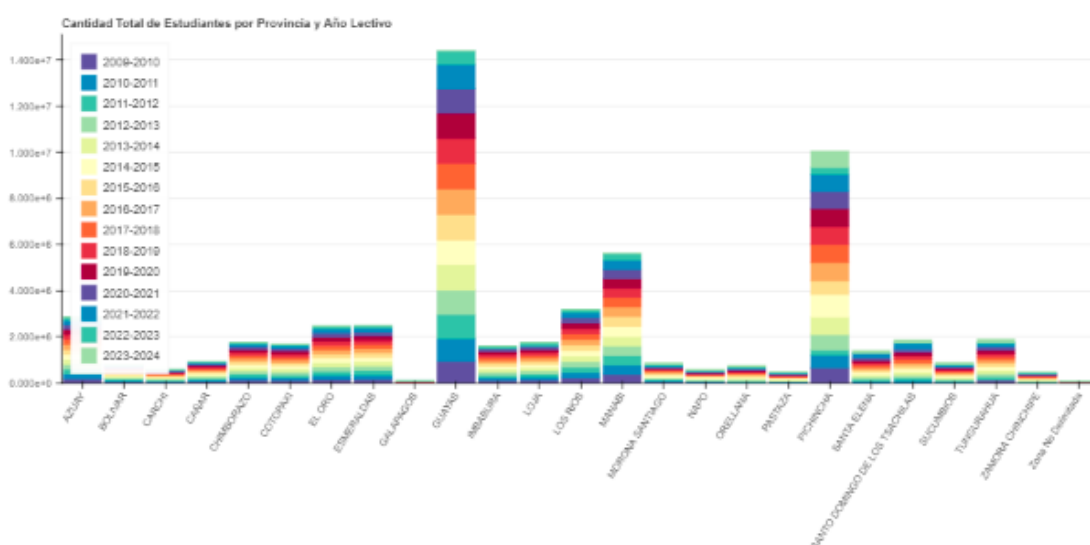


Figura 3: Total de estudiantes por provincia y año lectivo

#### Hallazgos:

- Como vimos también en el gráfico de barras de Matplotlib, algunas provincias, como Guayas y Pichincha, tienen números significativamente mayores de estudiantes en comparación con otras provincias.
- La variación en los colores a lo largo de los años dentro de cada barra provincial muestra cómo ha cambiado la cantidad de estudiantes a lo largo del tiempo. Por ejemplo, se puede observar que en ciertas provincias hay años con picos significativos o caídas.
- Este tipo de gráfico permitió comparar entre provincias en cantidades totales de estudiantes y en la evolución de esas cantidades con los años.

**Gráficos de barras (Bokeh):** compara las cantidades de estudiantes en diferentes categorías de resultados académicos a través de las provincias. El eje X (Horizontal) enumera las provincias. Cada segmento de la barra en el gráfico corresponde a una provincia específica. El eje Y (Vertical) representa la cantidad de estudiantes. También se especifican colores de las barras: verde - estudiantes promovidos, rojo - estudiantes no promovidos y azul - estudiantes que abandonaron.

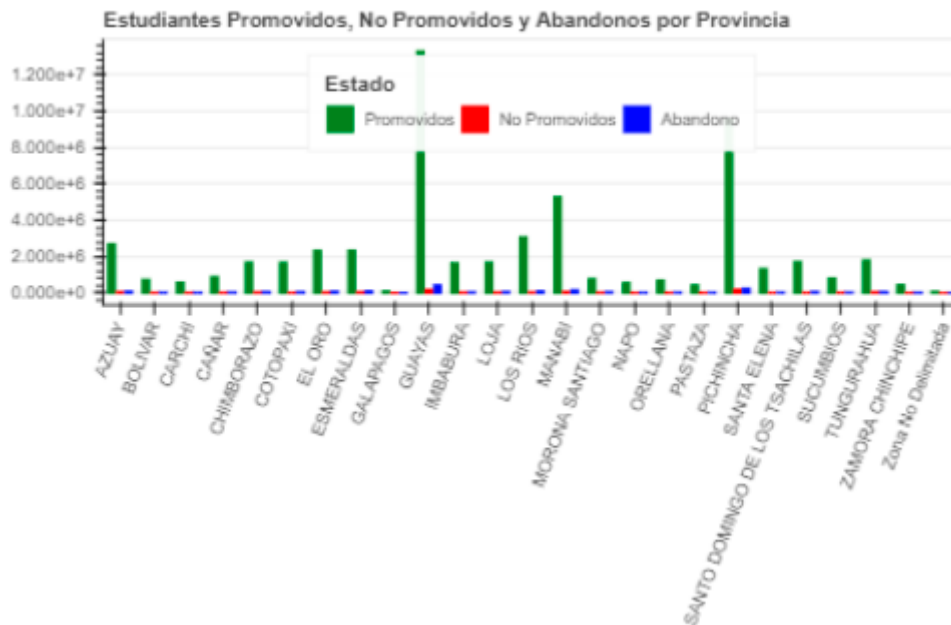


Figura 4: promovidos, no promovidos y abandonos por provincia

#### Hallazgos:

- La distribución de estudiantes en las tres categorías es notablemente desigual entre las provincias. Algunas provincias, como Guayas y Pichincha, muestran volúmenes significativamente mayores en comparación con otras. Que también lo vimos en los gráficos anteriores.
- En la mayoría de las provincias, el número de estudiantes promovidos es sustancialmente mayor que los no promovidos y los que abandonan, indicado por las barras verdes que se visualizan en el gráfico.
- En ciertas provincias, hay picos notables de abandono, lo que podría indicar problemas específicos en el entorno educativo, socioeconómico de esas áreas u otras variables a analizar.

**Gráfico de barras (Pywalker):** representa la cantidad total de estudiantes que abandonaron sus estudios, organizada por año lectivo desde 2010-2011 hasta 2023-2024. Esta visualización nos ayuda a identificar tendencias en el tiempo respecto al abandono escolar. El eje X (Horizontal) muestra los años lectivos, indicados en un formato que va desde el comienzo de una década escolar hasta la siguiente (2010-2011, 2011-2012, etc.). El eje Y (Vertical) representa la suma de estudiantes que abandonaron, escalado de 0 a 200,000 estudiantes.

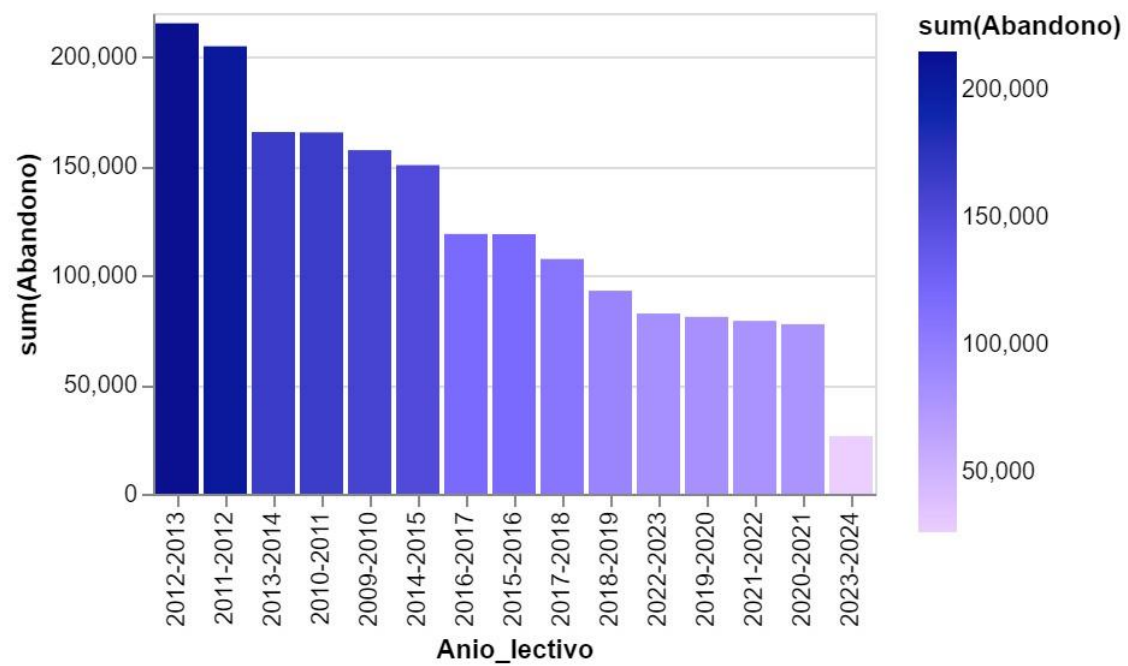


Figura 5: Abandono por año lectivo

Hallazgos:

- Hay una tendencia general a la baja en el número de abandonos a lo largo de los años, iniciando con un pico en 2010-2011.
- Entre 2010-2011 y 2014-2015, la disminución en el número de abandonos es bastante pronunciada, lo cual podría sugerir la implementación de cambios significativos en el entorno educativo que favorecen la retención escolar.
- Después de una relativa estabilización en el abandono escolar desde 2015 hasta 2020, la tendencia vuelve a caer a partir de 2020-2021.

**Gráfico de barras (Pywalker):** muestra la suma de estudiantes promovidos por provincia. Se utilizó para comparar cantidades absolutas entre varias categorías, en este caso, las provincias. El eje X (Horizontal) representa las diferentes provincias y El eje Y (Vertical) indica la suma total de estudiantes promovidos.



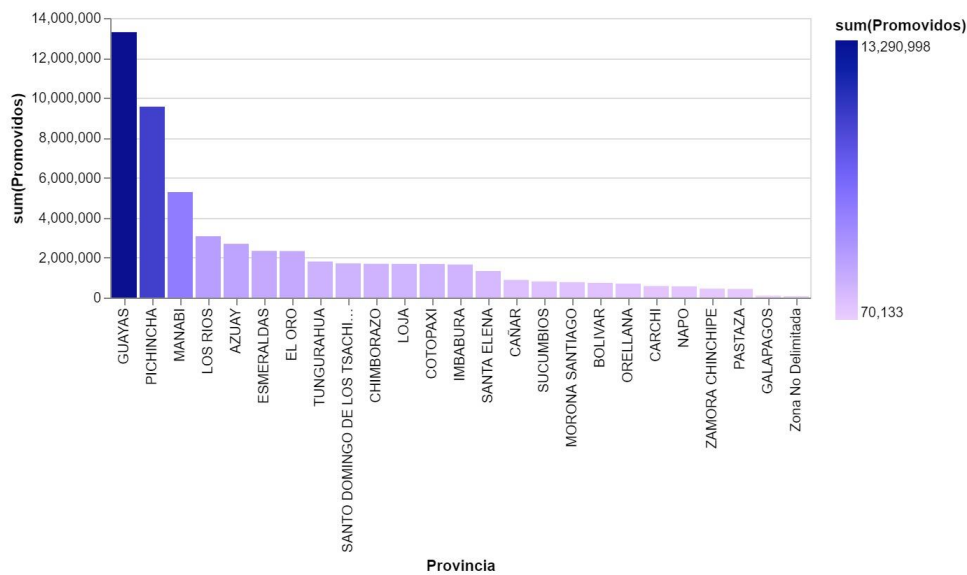


Figura 6: promovidos por provincia

#### Hallazgos:

- Existe una notable diferencia en el número de estudiantes promovidos entre las provincias. Guayas y Pichincha muestran cantidades significativamente más altas en comparación con otras provincias.
- Las provincias con grandes centros urbanos (Guayas y Pichincha) tienen un número mucho mayor de estudiantes promovidos, se podría asumir que refleja mayores poblaciones y posiblemente mejor acceso a recursos educativos.
- Provincias como Galápagos y Zamora Chinchipe muestran las cantidades más bajas de estudiantes promovidos, lo que podría indicar menores poblaciones estudiantiles o desafíos particulares en estos entornos.

## 5. CONCLUSIONES

- La utilización de herramientas de IA como pandas para el procesamiento de datos y Matplotlib, Bokeh y Pywalker para la visualización ha demostrado ser eficiente en el manejo de grandes volúmenes de datos educativos que tenemos en los data sets. Este tipo de herramientas nos pueden facilitar la identificación de tendencias clave y la toma de decisiones basada en datos. A través de estas herramientas hemos podido para esta aplicación, extraer información sobre promociones, abandonos y la distribución de estudiantes por provincias y años.
- El análisis que se realizó mostró varios hallazgos como se muestran en cada gráfica, así hubo variabilidad significativa en los patrones educativos entre provincias.

## 6. BIBLIOGRAFÍA

McKinney, W. (2012). Python for data analysis. O'Reilly Media, Inc.

Castro Quintero, N. J. (2017). Introducción a data science con Python

Alberola Oltra, J. M. (2021). Matplotlib.

Esnaola, L. (2022). Experiencia de uso de Google Colaboratory en Sistemas Inteligentes.

Elizalde Solano, R. R. (2024). Herramientas para inteligencia artificial: Guía didáctica. Facultad de Ingenierías y Arquitectura, Universidad Técnica Particular de Loja.

## 7. ANEXOS

## DICCIONARIO DE DATOS:

### Registro histórico 1.csv

Variable	Tipo de variable	Descripción
Anio_lectivo	Alfanumérica	Periodo escolar codificado para registros inicio o fin.  Inicio: Levantamiento cuyo objetivo es obtener información estadística al empezar el periodo escolar, el cual refleja la información de matrícula. Fin: Levantamiento cuyo objetivo es obtener información estadística al culminar el periodo escolar, el cual determina la descomposición de la matrícula (Promovidos, No promovidos, Abandono).
Zona	Alfanumérica	Zona de planificación, según ubicación geográfica de la institución educativa
Provincia	Alfabética	Corresponde al nombre de la provincia en la que se encuentra ubicada la institución educativa.
Cod_Provincia	Alfanumérica	Establece el código provincial, según la división política administrativa de la ubicación de la institución educativa.
Canton	Alfanumérica	Corresponde al nombre del cantón en la que se encuentra ubicada la institución educativa.
Cod_Canton	Alfanumérica	Establece el código cantonal, según la división política administrativa de la ubicación de la institución educativa.
Parroquia	Alfanumérica	Corresponde al nombre de la parroquia en la que se encuentra ubicada la institución educativa.
Cod_Parroquia	Alfanumérica	Establece el código parroquial, según la división política administrativa de la ubicación de la institución educativa.
Nombre_Institucion	Alfanumérica	Detalla la denominación de la institución educativa conforme el permiso de funcionamiento otorgado.
AMIE	Alfanumérica	Identifica de manera única a cada institución educativa, conformada por 8 caracteres.
Tipo_Educacion	Alfabética	Hace referencia al servicio educativo que oferta la institución conforme el grupo objetivo:  <b>Educación Ordinaria.</b> - Consiste en la oferta para estudiantes que asisten regularmente a clases y están en la edad en correspondiente al curso o grado (conforme lo establece el artículo 27 del Reglamento General a la LOEI). <b>Educación Especial.</b> - Considera en la oferta para los estudiantes con necesidades educativas especiales asociadas a una discapacidad. <b>Popular Permanente.</b> - Corresponde al servicio educativo para las y los jóvenes y adultos que no han podido culminar su escolarización en la edad correspondiente.

Sostenimiento	Alfabética	<p>Clasifica a la institución educativa conforme la proveniencia de los recursos que obtiene para desarrollar sus actividades correctamente, esta contiene las siguientes categorías:</p> <p><b>Fiscal.-</b> Se encuentran las instituciones cuya fuente de financiamiento proviene del Gobierno Central.</p> <p><b>Fiscomisional.-</b> Se encuentran las instituciones de derecho privado con apoyo estatal que brindan un servicio educativo complementario al del Estado.</p> <p><b>Municipal.-</b> Se encuentran las instituciones cuya fuente de financiamiento proviene del Gobierno Autónomos Descentralizados.</p> <p><b>Particular.-</b> Se encuentran las instituciones educativas constituidas y administradas por personas naturales o jurídicas de derecho privado.</p>
area	Alfabética	Corresponde a la ubicación geográfica urbana o rural de la institución educativa, conforme su parroquia.
Regimen_Escolar	Alfabética	Permite identificar el cronograma escolar que rige en la institución educativa en Costa - Galápagos o Sierra - Amazonía.
Jurisdiccion	Alfabética	Identifica a las instituciones educativas conforme su administración en Interculturales Bilingüe e Interculturales.
Docentes_Femenino	Numérica	Número de docentes por sexo femenino, registrados por la institución educativa
Docentes_Masculino	Numérica	Número de docentes por sexo masculino, registrados por la institución educativa
Total_Docentes	Numérica	Número total de docentes del Sistema Educativo Nacional, registrados por la institución educativa
Estudiantes_Femenino	Numérica	Número de estudiantes por sexo femenino, registrados por la institución educativa.
Estudiantes_Masculino	Numérica	Número de estudiantes por sexo masculino, registrados por la institución educativa.
Total_Estudiantes	Numérica	Número total de estudiantes del Sistema Educativo Nacional, registrados por la institución educativa.
Ecuatoriana	Numérica	País de procedencia del estudiante, registrados por la institución educativa
Colombiana	Numérica	País de procedencia del estudiante, registrados por la institución educativa
Venezolana	Numérica	País de procedencia del estudiante, registrados por la institución educativa
Peruana	Numérica	País de procedencia del estudiante, registrados por la institución educativa
Otros_Paises_de_America	Numérica	País de procedencia del estudiante, registrados por la institución educativa
Otros_Continentes	Numérica	País de procedencia del estudiante, registrados por la institución educativa

## Registro histórico 2.csv

Anio_lectivo	Alfanumérica	Periodo escolar codificado para registros inicio o fin. Inicio: Levantamiento cuyo objetivo es obtener información estadística al empezar el periodo escolar, el cual refleja la información de matrícula. Fin: Levantamiento cuyo objetivo es obtener información estadística al culminar el periodo escolar, el cual determina la descomposición de la matrícula (Promovidos, No promovidos, Abandono).
Zona	Alfanumérica	Zona de planificación, según ubicación geográfica de la institución educativa
Provincia	Alfabética	Corresponde al nombre de la provincia en la que se encuentra ubicada la institución educativa.
Cod_Provincia	Alfanumérica	Establece el código provincial, según la división política administrativa de la ubicación de la institución educativa.
Canton	Alfanumérica	Corresponde al nombre del cantón en la que se encuentra ubicada la institución educativa.
Cod_Canton	Alfanumérica	Establece el código cantonal, según la división política administrativa de la ubicación de la institución educativa.
Parroquia	Alfanumérica	Corresponde al nombre de la parroquia en la que se encuentra ubicada la institución educativa.
Cod_Parroquia	Alfanumérica	Establece el código parroquial, según la división política administrativa de la ubicación de la institución educativa.
Nombre_Institucion	Alfanumérica	Detalla la denominación de la institución educativa conforme el permiso de funcionamiento otorgado.
AMIE	Alfanumérica	Identifica de manera única a cada institución educativa, conformada por 8 caracteres.
Escolarizacion	Alfabética	La educación escolarizada conduce a la obtención de los siguientes títulos y certificados: el certificado de asistencia a la Educación Inicial, el certificado de terminación de la Educación General Básica y el título de Bachillerato.
Tipo_Educacion	Alfabética	Hace referencia al servicio educativo que oferta la institución conforme el grupo objetivo:  <b>Educación Ordinaria.</b> - Consiste en la oferta para estudiantes que asisten regularmente a clases y están en la edad en correspondiente al curso o grado (conforme lo establece el artículo 27 del Reglamento General a la LOEI). <b>Educación Especial.</b> - Considera en la oferta para los estudiantes con necesidades educativas especiales asociadas a una discapacidad. <b>Popular Permanente.</b> - Corresponde al servicio educativo para las y los jóvenes y adultos que no han podido culminar su escolarización en la edad correspondiente.

Sostenimiento	Alfabética	<p>Clasifica a la institución educativa, conforme la proveniencia de los recursos que obtiene para desarrollar sus actividades correctamente, esta contiene las siguientes categorías:</p> <p><b>Fiscal.-</b> Se encuentran las instituciones cuya fuente de financiamiento proviene del Gobierno Central.</p> <p><b>Fiscomisional.-</b> Se encuentran las instituciones de derecho privado con apoyo estatal que brindan un servicio educativo complementario al del Estado.</p> <p><b>Municipal.-</b> Se encuentran las instituciones cuya fuente de financiamiento proviene del Gobierno Autónomos Descentralizados.</p> <p><b>Particular.-</b> Se encuentran las instituciones educativas constituidas y administradas por personas naturales o jurídicas de derecho privado.</p>
area	Alfabética	Corresponde a la ubicación geográfica urbana o rural de la institución educativa, conforme su parroquia.
Regimen_Escolar	Alfabética	Permite identificar el cronograma escolar que rige en la institución educativa en Costa - Galápagos o Sierra - Amazonía.
Jurisdiccion	Alfabética	Identifica a las instituciones educativas conforme su administración en Interculturales Bilingüe e Interculturales.
Modalidad	Alfabética	<p>Establece el tipo de dinámica de trabajo mediante el cual se realiza el proceso de enseñanza y aprendizaje, determinado por las estrategias y formas de comunicación empleadas. La institución puede ofertar una o varias modalidades.</p> <p><b>Presencial.</b> Corresponde a la oferta que exige la asistencia regular de los estudiantes al establecimiento educativo</p> <p><b>Semipresencial.</b> No exige la asistencia regular de los estudiantes a los establecimientos educativos, y requiere el trabajo independiente de los estudiantes con un acompañamiento de la institución presencial periódica.</p> <p><b>A distancia.</b> Promueve el trabajo autónomo de los estudiantes con un acompañamiento no presencial de la institución educativa.</p> <p><b>Educación Abierta.</b> Desarrolla el proceso de enseñanza-aprendizaje sin exigir la asistencia regular de los estudiantes a la institución educativa.</p> <p><b>Educación en Casa.</b> Se basa en la educación en el cual padres, madres de familia o representantes del estudiante asumen la responsabilidad de educar directamente, o a través de docentes-tutores de Educación en Casa, a sus hijos e hijas o representados, con el apoyo y seguimiento de una Institución Educativa de cualquier sostenimiento, que regularice y garantice su derecho a la educación.</p>

Jornada	Alfabética	<p>Establece la proporción del día en que los estudiantes y docentes interactúan en el proceso de enseñanza aprendizaje (la institución educativa puede ofertar una o más de una jornada en el mismo periodo (Matutina, Vespertina, Nocturna)</p> <p><b>Matutina.-</b> Actividades que se desarrolla en la mañana</p> <p><b>Vespertina.-</b> Actividades que se desarrolla en la tarde</p> <p><b>Nocturna.-</b> Actividades que se desarrolla en la noche</p>
Acceso_Edificio	Alfabética	<p>Establece como las personas que no integran a la comunidad pueden acceder a la institución educativa, en las categorías:</p> <p><b>Terrestre.-</b> Acceso mediante calle, carretera, camino, herradura, sendero y verano</p> <p><b>Aérea.-</b> Acceso mediante avión, avioneta</p> <p><b>Fluvial.-</b> Acceso mediante barco, gabarra</p>
Total_Estudiantes	Numérica	Número total de estudiantes del Sistema Educativo Nacional, registrados por la institución educativa.
Promovidos	Numérica	Total, de estudiantes que al finalizar el año lectivo cumplieron los requisitos para continuar al siguiente grado.
No promovidos	Numérica	Total, de estudiantes que al finalizar el año lectivo no cumplieron los requisitos para continuar al siguiente grado.
Abandono	Numérica	Total, de estudiantes que se registraron en el año lectivo abandonaron los estudios durante el mismo año lectivo.