Universidad Central del Ecuador

FAC U LTAD D E I N G E N I E R ÍA Y

C I E N C IAS APLI CAD AS

**Sistemas De Información**

Documento de Análisis de Datos



**Estudiante:**

* Luis Angel Gaona Cumbicus

[lagaona@uce.edu.ec](mailto:lagaona@uce.edu.ec)

* Raul Alexander Pazos Erraez

[rapazos@uce.edu.ec](mailto:rapazos@uce.edu.ec)

* Cristian Daniel Toca Rocha

[cdtoca@uce.edu.ec](mailto:cdtoca@uce.edu.ec)

* Marlon Josue Espinosa Mancero

mjespinosam@uce.edu.ec

**Docente:**

* PhD, Jefferson Tarcisio Beltrán Morales

[jtbeltran@uce.edu.ec](mailto:jtbeltran@uce.edu.ec)

**Asignatura:** Minería de datos

**Paralelo:** S8-P2

**Fecha:** martes 2 de julio de 2024

Logotipo

Descripción generada automáticamente

**Generative AI-Powered Economic Impact Analysis System (GEIA)**

***Fecha:****28/06/2024*

Contenido

[HOJA DE CONTROL 4](#_Toc170811558)

[Historial de Cambios 4](#_Toc170811559)

[Introducción 5](#_Toc170811560)

HOJA DE CONTROL

|  |  |  |  |
| --- | --- | --- | --- |
| Organismo | Universidad Central Del Ecuador | | |
| Proyecto | Generative AI-Powered Economic Impact Analysis System (**GEIA**) | | |
| Entregable | Documento de análisis de los datos | | |
| Autor | Luis Angel Gaona Cumbicus | | |
| Versión/Edición | V1.0 | Fecha Versión | 02/07/2024 |
| Aprobado por |  | Fecha Aprobación | .…./…../….. |
|  |  | N.º Total de Páginas | 0 |

Historial de Cambios

|  |  |  |  |
| --- | --- | --- | --- |
| Fecha | Autor | Organización | Descripción |
|  |  |  |  |

Introducción

Este documento describe el análisis de los datos para el proyecto **“GEIA”** que consisten en desarrollar un modelo predictivo para estimar los daños económicos en diferentes sectores de Ecuador causados por eventos adversos. Este documento abarca el análisis exploratorio de los datos históricos utilizados para entrenar el modelo predictivo. Se describen las herramientas utilizadas, las variables analizadas y la preparación inicial de los datos.

Herramienta

Para el análisis de datos del proyecto se utilizó la herramienta **DiscoveR**. Esta herramienta ofrece un entorno intuitivo y potente para la exploración de datos, la preparación de datos y la visualización de resultados, lo que la convierte en una herramienta ideal para este tipo de análisis.

Descripción de las variables

Las variables utilizadas en el análisis se detallan a continuación:

* Año -Trimestre: Representa el año y trimestre en que se registró de los datos del VAB y los EVENTOS. Los trimestres están estandarizados de la siguiente manera:
  + **T1:** Enero, febrero y marzo
  + **T2:** Abril, mayo y junio
  + **T3:** Julio, agosto y septiembre
  + **T4:** Octubre, noviembre y diciembre **(Tiempo)**
* **Industria:** Variable categórica que representa el nombre del sector económico. Se consideran los 20 sectores principales de la economía ecuatoriana, como: "Agricultura ganadería y silvicultura", "Pesca y acuicultura", "Explotación de minas y canteras", "Manufactura de productos alimenticios", "Manufactura de productos no alimenticios", "Refinados de petroleo", "Suministro de electricidad y agua", "Construcción", "Comercio", "Transporte y almacenamiento", "Alojamiento y comidas", "Información y comunicación", "Actividades financieras y de seguros", "Actividades inmobiliarias", "Actividades profesionales técnicas", "Administración pública", "Enseñanza", "Salud y asistencia social", "Arte entretenimiento y otras actividades de servicios", "Actividades de los Hogares como empleadores". **(Categórica)**
* Evento: Variable categórica que describe el tipo de evento adverso que se produjo. Ejemplos de eventos incluyen: "Terremotos", "Inundaciones", "Deslizamientos de tierra", "Erupciones volcánicas", "Sequías", "Protestas", "manifestaciones", "Huelgas", "Actos de violencia", "Movimientos sociales", "Elecciones", "Cambios de gobierno", "Acuerdos y tratados internacionales", "Crisis políticas", "Relaciones diplomáticas", "Procesos electorales y reformas políticas", "Crisis económicas", "inflación", "desempleo", "Deuda pública", "Pandemias", "Epidemias", "Terrorismo", "Crisis energéticas", "Paros nacionales", "Temblores", "Eventos culturales y deportivos". **(Categórica)**
* **VAB:** Representa el Valor Agregado Bruto (VAB) del sector en miles de dólares **($)**. El VAB es un indicador económico que refleja el valor total de la producción de un sector en un período determinado. **(Numérica)**

Análisis de datos

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Interfaz de usuario gráfica, Tabla

Descripción generada automáticamenteInterfaz de usuario gráfica, Tabla

Descripción generada automáticamenteInterfaz de usuario gráfica, Tabla

Descripción generada automáticamenteInterfaz de usuario gráfica, Tabla

Descripción generada automáticamenteInterfaz de usuario gráfica, Tabla

Descripción generada automáticamenteInterfaz de usuario gráfica, Tabla

Descripción generada automáticamente

Este análisis de datos sobre el Valor Agregado Bruto (VAB) en diferentes sectores de la economía ecuatoriana muestra la variabilidad y el desempeño económico en áreas clave como agricultura, pesca, minería, manufactura, comercio, entre otros. Se observan estadísticas descriptivas como el promedio, la mediana y la desviación estándar para cada sector, destacando diferencias significativas entre sectores. Además, se identifican eventos importantes como la pandemia de COVID-19, crisis económicas, terremotos, y otros, que han impactado en estos sectores, sugiriendo correlaciones entre los eventos y las fluctuaciones en el VAB.

* **Agricultura, ganadería y silvicultura**: Este sector muestra una amplia variabilidad con un VAB que va desde 441,331.1 hasta 1,965,091. La desviación estándar es relativamente alta, lo que indica fluctuaciones significativas en su contribución económica a lo largo del tiempo.
* **Pesca y acuicultura:** Aunque más estable que la agricultura, también presenta variaciones notables con un mínimo de 81,288.6 y un máximo de 435,385.2. Este sector parece ser menos volátil comparado con otros.
* **Explotación de minas y canteras:** Este sector es uno de los más importantes en términos de VAB, con un promedio superior a los 2 millones y un máximo de 4,485,167. Sin embargo, presenta una alta desviación estándar, indicando una considerable volatilidad.
* **Manufactura de productos alimenticios y no alimenticios:** Ambos sectores presentan una contribución económica considerable y estable, con valores promedio de alrededor de 1 millón y máximos que superan los 2 millones. La estabilidad relativa se refleja en sus desviaciones estándar.
* **Comercio:** Este sector es uno de los pilares de la economía ecuatoriana con un VAB que oscila entre 597,859 y 4,958,755. La alta desviación estándar sugiere que su desempeño está fuertemente influenciado por eventos externos.
* **Construcción:** Este sector muestra un rango amplio en su VAB, con un mínimo de 69,262.71 y un máximo de 1,879,541. La alta desviación estándar refuerza la idea de que la construcción está sujeta a fluctuaciones significativas, posiblemente vinculadas a inversiones públicas y privadas.
* **Eventos Significativos:** Se observan eventos recurrentes que afectan la economía, como inundaciones, protestas y crisis políticas. La pandemia de COVID-19 y crisis económicas también tienen un impacto considerable, reflejado en la variabilidad del VAB en diversos sectores. Estos eventos pueden causar descensos o cambios bruscos en la economía.
* **VAB Total y PIB:** El VAB total muestra una considerable fluctuación, con un mínimo de 3,403,415 y un máximo de 28,250,684. Esto está alineado con las variaciones del PIB, que también tiene un rango amplio y una alta desviación estándar.

Gráfico

Descripción generada automáticamenteGráfico

Descripción generada automáticamenteGráfico, Gráfico de líneas

Descripción generada automáticamenteCaptura de pantalla de computadora

Descripción generada automáticamente

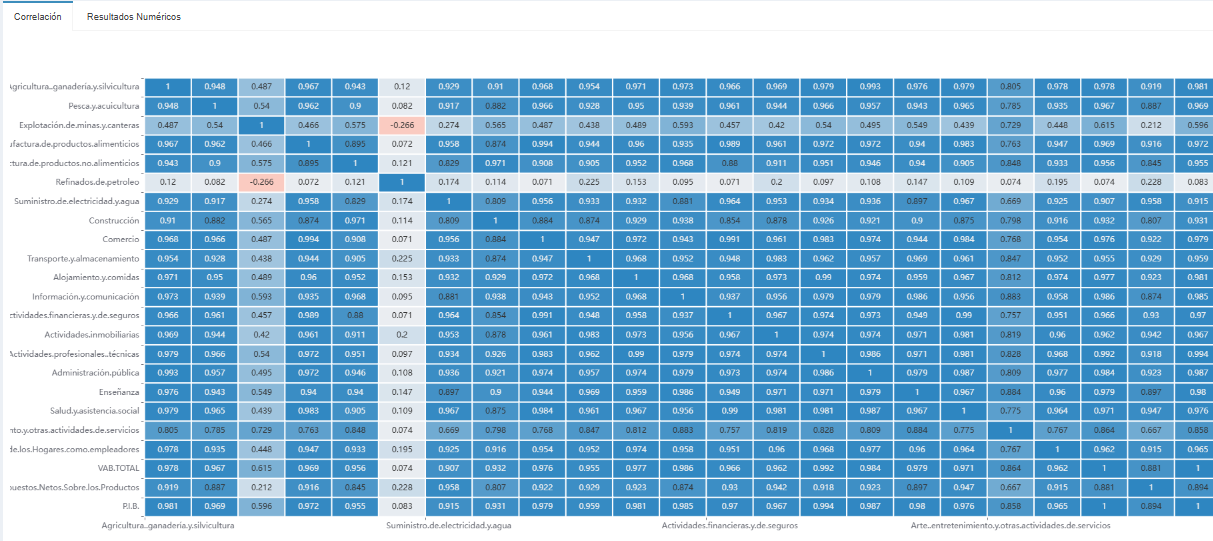
Una observación clave es que las variables analizadas no presentan una distribución normal. Esto se evidencia por la alta variabilidad y la amplitud de los rangos en los datos de Valor Agregado Bruto (VAB) en diversos sectores. La presencia de grandes desviaciones estándar y la amplia diferencia entre los valores mínimo y máximo sugieren que los datos están sesgados o tienen una distribución asimétrica, lo que podría indicar la influencia de eventos atípicos o fluctuaciones significativas en la economía ecuatoriana. Esto implica que los modelos estadísticos basados en la normalidad pueden no ser apropiados sin una transformación previa de los datos.

Gráfico, Gráfico de dispersión

Descripción generada automáticamente Gráfico, Gráfico de dispersión

Descripción generada automáticamente Gráfico

Descripción generada automáticamente Gráfico, Gráfico de barras

Descripción generada automáticamente  Tabla

Descripción generada automáticamente

El análisis de correlación entre las variables revela que no todas siguen una distribución normal. Esto se evidencia en las fuertes correlaciones observadas entre ciertos sectores, como **Agricultura, Ganadería y Silvicultura** con **Comercio** (correlación de 0.968), y **Manufactura de Productos Alimenticios** con **Actividades Financieras y de Seguros** (correlación de 0.988). Sin embargo, sectores como **Refinados de Petróleo** presentan correlaciones bajas o incluso negativas con otros sectores, como su correlación con **Explotación de Minas y Canteras** (-0.266). Estos resultados indican una diversidad en la relación entre las variables económicas, lo que sugiere que el comportamiento de algunas variables podría estar influenciado por factores específicos de cada sector o por eventos externos que afectan selectivamente a ciertas industrias. **Enseñanza** y **Administración pública** tienen una alta correlación con varias otras categorías, especialmente con **Información y comunicación** (0.98550.98550.9855 y 0.97890.97890.9789 respectivamente), y **Salud y asistencia social** (0.96650.96650.9665 y 0.98720.98720.9872 respectivamente). Esto sugiere que los sectores educativos y de administración pública están estrechamente relacionados con la actividad económica general y el bienestar social. **Actividades financieras y de seguros** y **Actividades inmobiliarias** tienen una alta correlación (0.96670.96670.9667). Ambos sectores tienden a estar vinculados a la salud económica general y pueden ser influenciados por las mismas variables macroeconómicas. La correlación entre **VAB Total** y **PIB** es extremadamente alta (0.99960.99960.9996), indicando que estos dos indicadores están prácticamente en sincronía. Esto refuerza la idea de que el valor agregado bruto total es una buena representación del PIB. **Manufactura de productos alimenticios** y **Manufactura de productos no alimenticios** muestran una fuerte correlación entre sí (0.93510.93510.9351 y 0.96800.96800.9680). De igual manera, **Agricultura, ganadería y silvicultura** también tienen una alta correlación con otros sectores relacionados con la producción (0.97300.97300.9730 con **Pesca y acuicultura**). **Refinados de petróleo** muestra la menor correlación con muchos otros sectores, lo que puede indicar que este sector tiene una dinámica económica distinta comparada con otros sectores más interrelacionados. **Arte, entretenimiento y otras actividades de servicios** presenta una correlación moderada con varios sectores (0.77540.77540.7754 a 0.88380.88380.8838), mostrando que aunque tiene cierta relación con la economía en general, su influencia es menos directa en comparación con sectores como **Construcción** y **Comercio**. La correlación de **Impuestos netos sobre los productos** con **PIB** es moderada (0.89370.89370.8937), indicando que los impuestos también afectan significativamente el PIB, aunque no tanto como algunos otros factores. **Salud y asistencia social** muestra una alta correlación con **Actividades profesionales, técnicas** (0.98120.98120.9812), **Administración pública** (0.98720.98720.9872), y **Enseñanza** (0.96650.96650.9665). Esto puede reflejar la interdependencia entre el bienestar social y las actividades relacionadas con la administración y la educación.

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

Diagrama

Descripción generada automáticamente Gráfico, Gráfico de dispersión

Descripción generada automáticamente Tabla

Descripción generada automáticamente

Imagen que contiene Gráfico

Descripción generada automáticamente

Gráfico, Gráfico de dispersión

Descripción generada automáticamente Interfaz de usuario gráfica, Gráfico

Descripción generada automáticamente con confianza media Gráfico, Gráfico radial

Descripción generada automáticamente Gráfico, Gráfico de barras

Descripción generada automáticamente

**Resumen de los Clusters**

**Cluster 1:**

* **Características Generales:**
  + El Cluster 1 está compuesto principalmente por los datos de los primeros años (2000 a 2012).
  + Este cluster muestra valores negativos en la mayoría de los indicadores económicos.
* **Características Económicas:**
  + **Agricultura, Ganadería y Silvicultura**: Valores negativos, lo que indica que estos sectores tienen un menor desempeño en comparación con el promedio.
  + **Pesca y Acuicultura**: También negativo.
  + **Explotación de Minas y Canteras**: Negativo.
  + **Manufactura de Productos Alimenticios y No Alimenticios**: Negativos.
  + **Construcción y Comercio**: Negativos.
  + **Transporte y Almacenamiento, Alojamiento y Comidas, Información y Comunicación**: Valores negativos.
  + **Salud y Asistencia Social**: Negativo.
  + **VAB Total e Impuestos Netos Sobre los Productos**: Negativos.
  + **PIB**: Negativo.

**Cluster 2:**

* **Características Generales:**
  + El Cluster 2 abarca desde 2013 hasta 2023.
  + Muestra valores positivos en la mayoría de los indicadores económicos, indicando un mejor desempeño en comparación con el promedio.
* **Características Económicas:**
  + **Agricultura, Ganadería y Silvicultura**: Valores positivos.
  + **Pesca y Acuicultura**: Positivos.
  + **Explotación de Minas y Canteras**: Positivos.
  + **Manufactura de Productos Alimenticios y No Alimenticios**: Positivos.
  + **Refinados de Petróleo y Suministro de Electricidad y Agua**: Positivos.
  + **Construcción y Comercio**: Positivos.
  + **Transporte y Almacenamiento, Alojamiento y Comidas, Información y Comunicación**: Positivos.
  + **Actividades Financieras y de Seguros, Actividades Inmobiliarias**: Positivos.
  + **Actividades Profesionales, Técnicas, Administración Pública, Enseñanza, Salud y Asistencia Social**: Positivos.
  + **Arte, Entretenimiento y Otras Actividades de Servicios**: Positivos.
  + **Actividades de los Hogares como Empleadores, VAB Total, Impuestos Netos Sobre los Productos**: Positivos.
  + **PIB**: Positivo.

**Interpretación y Recomendaciones**

1. **Evolución Temporal:**
   * Los clusters muestran una clara diferencia en el desempeño económico entre los periodos anteriores a 2013 y los posteriores. Esto sugiere un cambio significativo en la estructura económica o en los indicadores medidos.
2. **Características de los Clusters:**
   * **Cluster 1** representa un periodo con menor actividad económica en comparación con el promedio. Los sectores como Agricultura, Construcción y Comercio tuvieron un bajo desempeño.
   * **Cluster 2** representa un periodo con un crecimiento o mejora en los indicadores económicos, con todos los sectores mostrando valores positivos.
3. **Análisis de Variables:**
   * La diferencia en los centros de los clusters refleja que el Cluster 1 está compuesto por años y sectores económicos que experimentaron dificultades o menos crecimiento, mientras que el Cluster 2 refleja años y sectores que mostraron un crecimiento robusto.
4. **Aplicaciones Prácticas:**
   * **Para Políticas Económicas:** Las políticas deben adaptarse a los diferentes periodos identificados. El Cluster 2, con sus resultados positivos, puede ser un punto de referencia para estrategias futuras.
   * **Para Inversión:** Invertir en los sectores destacados en el Cluster 2 podría ser más prometedor, dado su desempeño superior.

**Estos insights te permitirán entender mejor las tendencias económicas a lo largo del tiempo y cómo los diferentes sectores se comportaron en los diferentes periodos.**