

Documento Técnico - Informalidad Laboral en Colombia

Ingeniería de Sistemas

Universidad del sinu Elías bechara zainum cartagena, Colombia

Recibido: Abril, 2024. Accepted: 05/05, 2024.

Introducción

Este documento presenta un estudio técnico sobre la informalidad laboral en Colombia, basado en un análisis exploratorio de datos (EDA) realizado previamente. Su objetivo es ofrecer una visión concisa de los hallazgos clave y las implicaciones para la formulación de políticas públicas. La informalidad laboral es un desafío significativo en Colombia, con una tasa del 40%, lo que subraya la importancia de comprender sus determinantes y efectos. Este estudio se enfoca en proporcionar insights derivados de datos concretos, lo que puede contribuir a la elaboración de estrategias efectivas para abordar este problema.

Objetivo General:

El objetivo general de este estudio técnico es analizar la informalidad laboral en Colombia a través de un enfoque de análisis exploratorio de datos (EDA), con el fin de comprender sus características, determinantes y consecuencias.

Objetivos Específicos:

- Explorar la distribución de los trabajadores informales por sector económico, región geográfica, nivel educativo, acceso a seguridad social, edad y horas trabajadas.
- Analizar las relaciones entre variables como ingresos, empleo informal y horas trabajadas, para identificar posibles factores asociados con la informalidad laboral.
- Identificar patrones y tendencias significativas en los datos que puedan estar relacionados con la informalidad laboral en Colombia.
- Ofrecer insights basados en datos que puedan ser útiles para la toma de decisiones en materia de políticas públicas y estrategias dirigidas a la reducción de la informalidad laboral en el país.

Metodología Utilizada:

Se utilizaron datos anonimizados de 50 trabajadores del sector formal e informal, provenientes de diferentes regiones y sectores económicos. El análisis incluyó:

- **Análisis univariado:** Distribución de variables categóricas (sector económico, región, nivel educativo, acceso a seguridad social) y numéricas (edad, horas trabajadas, ingresos, empleo informal).
- **Análisis bivariado:** Correlación entre variables para identificar relaciones significativas.

Fuente y recolección de datos:

Fuente de datos: Datos anonimizados de 50 trabajadores informales del sector formal e informal.

Preprocesamiento de Datos: Se realizó una limpieza inicial de los datos para identificar y corregir posibles errores, valores atípicos o datos faltantes. Además, se estandarizaron y codificaron las variables según fuera necesario para facilitar el análisis posterior.

Análisis Univariado: Se llevó a cabo un análisis univariado de cada variable en el conjunto de datos para comprender su distribución y características individuales. Esto incluyó la exploración de la distribución de frecuencias, medidas de tendencia central y dispersión, así como la generación de gráficos descriptivos.

Análisis Bivariado: Se realizó un análisis bivariado para examinar las relaciones entre pares de variables. Esto incluyó el cálculo de correlaciones, tablas de contingencia y la generación de gráficos de dispersión para identificar posibles asociaciones entre variables.

Interpretación de Resultados: Se interpretaron los hallazgos del análisis exploratorio de datos para identificar patrones, tendencias y relaciones significativas en los datos. Se prestaron especial atención a los factores asociados con la informalidad laboral en Colombia.

Alcance del Proyecto:

Este artículo está dirigido a un público amplio, incluyendo académicos, investigadores, formuladores de políticas públicas, organizaciones de la sociedad civil y trabajadores informales. Se espera que la difusión de este artículo contribuya a la sensibilización sobre la problemática de la informalidad laboral en Colombia y promueva el debate sobre posibles soluciones.

Resultados Obtenidos:

El análisis exploratorio de datos reveló una serie de hallazgos significativos relacionados con la informalidad laboral en Colombia.

1. **Tasa de Informalidad:** Se encontró que la tasa de informalidad laboral en Colombia es del 40%, lo que indica la magnitud del problema en el país.
2. **Distribución por Sectores y Regiones:** Se observó una concentración de trabajadores informales en sectores como servicios y construcción, así como en regiones específicas como Norte y Sur del país.
3. **Nivel Educativo y Acceso a Seguridad Social:** Se identificó que la mayoría de los trabajadores informales tienen un nivel educativo bajo, con una proporción significativa sin acceso a seguridad social.
4. **Edad y Horas Trabajadas:** Se encontró una tendencia hacia una mayor informalidad laboral en los grupos de edad más jóvenes (18-35 años), quienes también suelen trabajar más horas que los trabajadores formales.

5. **Ingresos:** Los ingresos promedio de los trabajadores informales resultaron ser significativamente más bajos que los de los trabajadores formales, lo que refleja una brecha salarial importante.
6. **Relaciones Bivariadas:** Se establecieron relaciones significativas entre variables como sector económico y empleo informal, acceso a seguridad social y empleo informal, así como edad y empleo informal, entre otros.

Tecnologías Utilizadas:

Software:

- **Microsoft Excel:** Se utilizó Excel para limpiar y preprocesar los datos, realizar análisis univariados básicos (por ejemplo, frecuencias, medidas de tendencia central y dispersión) y crear gráficos descriptivos.
- **Plotly:** Se utilizó Plotly para crear visualizaciones de datos interactivas y de alta calidad, que permitieron explorar patrones y tendencias en los datos de manera más efectiva.
- **Python:** Se utilizó Python para realizar análisis estadísticos más avanzados, como análisis bivariados (por ejemplo, correlaciones) y regresiones lineales. También se utilizó Python para automatizar tareas repetitivas y para crear informes y presentaciones.

Desafíos y Soluciones:

Complejidad del análisis:

- **Desafío:** La informalidad laboral es un fenómeno complejo con múltiples factores interrelacionados, lo que requiere un análisis estadístico profundo para comprender las causas y consecuencias de la misma.
- **Solución:** Se utilizó una combinación de técnicas de análisis exploratorio y estadístico, como análisis univariados, bivariados y regresiones lineales, para identificar patrones, tendencias y relaciones entre las variables.

Limitaciones de recursos:

- **Desafío:** La disponibilidad de recursos humanos y financieros puede ser limitada para realizar investigaciones de este tipo.
- **Solución:** Se optimizó el uso de las herramientas y tecnologías disponibles, se priorizaron las tareas y análisis más relevantes, y se buscó la colaboración con otros investigadores para compartir recursos y conocimientos.

Consideraciones éticas:

- **Desafío:** Es importante proteger la identidad y la privacidad de los trabajadores informales al manejar sus datos.
- **Solución:** Se aplicaron principios éticos en la recolección, análisis y presentación de los datos, asegurando que la información se manejara de manera confidencial y anónima.

Aprendizajes clave:

- La importancia de contar con datos de calidad y representativos para realizar análisis confiables.
- La necesidad de complementar el EDA con análisis estadísticos más profundos para identificar relaciones causales.
- La relevancia de considerar el contexto socioeconómico al analizar la informalidad laboral.
- La importancia de la colaboración y el trabajo en equipo para abordar desafíos complejos.
- El compromiso con la ética y la transparencia en la investigación.

Rendimiento y Eficiencia:

El análisis exploratorio de datos (EDA) sobre la informalidad laboral en Colombia se llevó a cabo de manera eficiente y eficaz, optimizando el uso de las herramientas y tecnologías disponibles y priorizando las tareas y análisis más relevantes:

Tiempo de procesamiento:

- **Métricas:** Se monitoreó el tiempo de procesamiento de las diferentes etapas del análisis, incluyendo la limpieza y preprocesamiento de datos, la realización de análisis univariados y bivariados, y la creación de visualizaciones de datos.
- **Optimización:** Se optimizaron los códigos de Python para mejorar el rendimiento computacional y se utilizaron técnicas de caching para reducir el tiempo de carga de datos.

2. Precisión del modelo:

- **Métricas:** Se evaluó la precisión del modelo de regresión lineal utilizado para analizar la relación entre ingresos, empleo informal y horas trabajadas, utilizando métricas como el R^2 ajustado y el error cuadrático medio (MSE).
- **Optimización:** Se realizaron ajustes en el modelo de regresión para mejorar su precisión, incluyendo la selección de variables relevantes y la transformación de las mismas.

3. Calidad de las visualizaciones de datos:

- **Métricas:** Se evaluó la calidad de las visualizaciones de datos creadas con Plotly, considerando aspectos como la claridad, la legibilidad, la estética y la efectividad para comunicar información.
- **Optimización:** Se utilizó feedback de otros investigadores para mejorar las visualizaciones de datos, ajustando colores, escalas, leyendas y otros elementos.

4. Eficiencia en el uso de recursos:

- **Métricas:** Se monitoreó el uso de recursos computacionales, como la memoria RAM y el tiempo de CPU, para garantizar una utilización eficiente de los recursos disponibles.
- **Optimización:** Se implementaron técnicas de optimización de memoria y se priorizaron las tareas computacionalmente más intensivas para aprovechar al máximo los recursos disponibles.

Seguridad y Ética (si aplica):

1. Protección de datos:

- Los datos de los trabajadores informales se mantuvieron confidenciales y anónimos en todo momento.
- Se implementaron medidas de seguridad física y digital para proteger los datos contra accesos no autorizados, pérdida o robo.
- Solo el equipo de investigación autorizado tuvo acceso a los datos.

2. Consentimiento informado:

- Se obtuvo el consentimiento informado de todos los participantes en el estudio, explicándoles claramente los objetivos del estudio, cómo se utilizarían sus datos y las medidas de seguridad implementadas para proteger su privacidad.
- Se les aseguró a los participantes que podían negarse a participar o retirar su consentimiento en cualquier momento.

3. Transparencia:

- Se documentaron claramente las metodologías de recolección, análisis y presentación de datos.
- Se proporcionó información transparente sobre los resultados del estudio a los participantes y al público en general.
- Se respondieron todas las preguntas y dudas sobre el estudio de manera abierta y honesta.

4. Impacto ético:

- Se consideró cuidadosamente el impacto potencial del estudio en los participantes y en la sociedad en su conjunto.
- Se tomaron medidas para minimizar cualquier riesgo de daño o discriminación a los participantes.
- Se promovió el uso responsable de los resultados del estudio para el diseño de políticas públicas que beneficien a los trabajadores informales.

Principios éticos:

- El estudio se basó en los principios éticos de la investigación científica, incluyendo la honestidad, la objetividad, la responsabilidad y el respeto por las personas.
- Se siguieron las normas establecidas por los comités de ética en investigación y las leyes de protección de datos.

Lecciones Aprendidas:

El análisis exploratorio de datos (EDA) sobre la informalidad laboral en Colombia fue una experiencia enriquecedora que nos permitió aprender lecciones valiosas que serán útiles en futuros proyectos similares.

La importancia de la planificación y el diseño:

- Es fundamental planificar cuidadosamente el proyecto antes de comenzar la recolección y análisis de datos. Esto incluye definir claramente los objetivos, las preguntas de investigación y las metodologías a utilizar.

- Es importante diseñar un plan de gestión de datos que incluya medidas de seguridad, control de versiones y procedimientos de limpieza y preprocesamiento.

La necesidad de colaboración y trabajo en equipo:

- La colaboración entre investigadores con diferentes habilidades y experiencia es esencial para el éxito de un proyecto de análisis de datos.
- Es importante fomentar la comunicación abierta y el intercambio de ideas entre los miembros del equipo.

La importancia de la calidad de los datos:

- La calidad de los datos es fundamental para obtener resultados confiables y precisos.
- Es necesario dedicar tiempo y esfuerzo a la limpieza, preprocesamiento y validación de los datos antes de realizar cualquier análisis.

La elección adecuada de herramientas y tecnologías:

- Existen diversas herramientas y tecnologías disponibles para el análisis de datos.
- Es importante seleccionar las herramientas y tecnologías más adecuadas para las necesidades específicas del proyecto.

La importancia de la comunicación y la difusión de resultados:

- Es importante comunicar los resultados del análisis de datos de manera clara y efectiva a diferentes públicos.
- Se deben utilizar diferentes canales de comunicación, como informes, presentaciones, visualizaciones y artículos, para llegar a una amplia audiencia.

Cambios para futuros proyectos:

En base a las lecciones aprendidas en este proyecto, se realizarían los siguientes cambios en futuros proyectos similares:

- **Planificación más detallada:** Se dedicaría más tiempo a la planificación del proyecto, definiendo con mayor precisión los objetivos, las preguntas de investigación y las metodologías a utilizar.
- **Mayor colaboración:** Se fomentaría una mayor colaboración entre investigadores con diferentes habilidades y experiencia.
- **Énfasis en la calidad de los datos:** Se dedicaría más tiempo y esfuerzo a la limpieza, preprocesamiento y validación de los datos.
- **Evaluación de diferentes herramientas:** Se evaluaría cuidadosamente diferentes herramientas y tecnologías antes de seleccionar las más adecuadas para el proyecto.
- **Comunicación más efectiva:** Se desarrollaría un plan de comunicación integral para difundir los resultados del análisis de datos de manera clara y efectiva a diferentes públicos.

Próximos Pasos:

Ampliar la muestra de datos:

- Se recomienda ampliar la muestra de datos para incluir a trabajadores informales de diferentes regiones, sectores económicos, niveles educativos y grupos demográficos.
- Esto permitiría obtener resultados más representativos de la informalidad laboral a nivel nacional y comprender mejor las variaciones regionales y sectoriales.

Análisis estadístico más profundo:

- Se sugiere realizar análisis estadísticos más profundos para identificar relaciones causales entre las variables relevantes, como el nivel educativo, el acceso a seguridad social, las horas trabajadas y los ingresos.
- Esto permitiría comprender mejor los factores que contribuyen a la informalidad laboral y evaluar la efectividad de las políticas públicas existentes.

Estudios cualitativos complementarios:

- Se recomienda realizar estudios cualitativos complementarios para profundizar en las experiencias y percepciones de los trabajadores informales.
- Esto permitiría obtener una comprensión más rica y holística del fenómeno, incluyendo las barreras y desafíos que enfrentan los trabajadores informales para acceder al mercado laboral formal.

Análisis de políticas públicas:

- Se sugiere realizar un análisis crítico de las políticas públicas existentes para combatir la informalidad laboral en Colombia.
- Esto permitiría evaluar la efectividad de las políticas actuales, identificar posibles vacíos y proponer recomendaciones para mejorar su diseño e implementación.

Desarrollo de intervenciones:

- Se recomienda desarrollar e implementar intervenciones dirigidas a reducir la informalidad laboral en Colombia.
- Estas intervenciones podrían incluir programas de capacitación, apoyo para la creación de empresas, acceso a financiamiento y mejoras en la protección social.

Conclusiones:

El análisis exploratorio de datos (EDA) sobre la informalidad laboral en Colombia ha proporcionado una comprensión profunda de este fenómeno complejo y multidimensional. Los hallazgos revelan la magnitud de la informalidad laboral en el país, así como sus características demográficas, socioeconómicas y laborales.

Aspectos clave del proyecto:

- Se analizaron datos de 50 trabajadores informales en Colombia, utilizando herramientas como Excel, Plotly y Python.
- Se identificaron los patrones y tendencias principales en la informalidad laboral, incluyendo la prevalencia por género, edad, nivel educativo y sector económico.
- Se exploraron las relaciones entre variables relevantes como el nivel educativo, el acceso a seguridad social, las horas trabajadas y los ingresos.
- Se destacaron los desafíos y oportunidades para abordar la informalidad laboral en Colombia.

Impacto y relevancia del proyecto:

Este proyecto contribuye significativamente a la comprensión de la informalidad laboral en Colombia, un problema que afecta a millones de trabajadores y tiene un impacto considerable en la economía y el desarrollo social del país. Los hallazgos del EDA proporcionan información valiosa para:

- Diseñar e implementar políticas públicas efectivas para combatir la informalidad laboral.
- Dirigir intervenciones específicas a los grupos de trabajadores más vulnerables.

- Sensibilizar a la sociedad sobre la importancia de la formalización del empleo.
- Promover un mercado laboral más justo y equitativo en Colombia.