

Minería de datos II

REGLAS DE ASOCIACIÓN

ECE 2T 2024

Andrés Mauricio Avilés Camacho

INTRODUCCIÓN AL PROYECTO



Este trabajo aplica técnicas de minería de datos —específicamente **reglas de asociación**— sobre la Encuesta Continua de Empleo del segundo trimestre de 2024 elaborada por el Instituto Nacional de Estadística de Bolivia (INE). El objetivo principal es identificar perfiles sociodemográficos típicos asociados a distintos tipos de ocupación en el país.

Para ello, se utilizan variables clave como sexo, edad, nivel educativo, zona geográfica y ocupación principal, con el fin de descubrir patrones frecuentes y relaciones relevantes entre estas características.

OBJETIVOS

GENERAL

Identificar perfiles sociodemográficos asociados la formalidad de laboral en Bolivia mediante el análisis de reglas de asociación aplicadas a datos de la Encuesta Continua de Empleo (ECE) del segundo trimestre de 2024.

ESPECÍFICOS

- Preparar y transformar el microdato de la ECE 2T-2024 mediante técnicas de preprocesamiento en R, seleccionando variables clave presentes en el dataset.
- Aplicar el algoritmo Apriori para extraer reglas de asociación frecuentes que vinculen combinaciones de características personales con formas específicas de ocupación e indicadores de formalidad laboral.
- Evaluar las reglas generadas mediante métricas como soporte, confianza y lift, seleccionando aquellas que presenten mayor relevancia explicativa y consistencia estadística.

MARCO TEÓRICO

- Alta informalidad estructural en Bolivia ($\approx 70\%$), con baja protección social y precariedad (Hurtado, 2022).
- Causas del empleo informal:
 - **Incentivos perversos:** Alta tributación para PYMES (Herbas-Torrico, 2020)
 - **Desajuste educativo:** Formación no alineada a demandas laborales
 - **Falla estructural:** Sector formal no absorbe mano de obra (especialmente jóvenes)
- "Mujeres en informalidad ganan 40% menos que hombres" (Urquidi et al., 2021)
- "Cada año de educación técnica reduce informalidad juvenil en 8%" (Vargas, 2011)
- "El sector informal reduce el crecimiento potencial en 1.5% anual" (Hurtado, 2022 - Modelo DSGE).

MOTIVACIÓN

El empleo informal en Bolivia afecta a gran parte de la población activa y está ligado a condiciones de vulnerabilidad. Pese a su importancia, las políticas públicas carecen de herramientas analíticas para identificar con precisión a los grupos más expuestos.

Este estudio propone usar reglas de asociación sobre datos de la ECE del INE para detectar perfiles típicos de informalidad, integrando variables como ocupación, edad, educación y tenencia de NIT.

METODOLOGÍA

ALGORITMO APRIORI

01

Variables relevantes

- s1_02 (sexo)
- s1_03a (edad)
- niv_ed_g (nivel educativo)
- area
- s2_18 (ocupación principal)
- s2_23 (NIT en el establecimiento de trabajo)

02

Procesamiento

- Se agruparon las edades en rangos para facilitar la interpretación.
- Se eliminaron registros con valores faltantes.
- Transformación del dataset a formato de transacciones

03

Extraer reglas

- Variables sociodemográficas que determinen el sector ocupacional.
- Variables sociodemográficas que determinen si el negocio donde trabajan tiene NIT.

RESULTADOS

DETERMINANTES DE TIPO DE OCUPACIÓN

Regla	Support	Confidence	Lift
Mujer, 15–24, rural, sin NIT → Familiar no remunerado	0.010	0.753	4.679
Mujer, 15–24, rural → Familiar no remunerado	0.010	0.734	4.559
Mujer, 15–24, superior, sin NIT → Familiar no remunerado	0.024	0.687	4.266
15–24, rural, superior, sin NIT → Familiar no remunerado	0.014	0.671	4.168
Mujer, 15–24, superior → Familiar no remunerado	0.025	0.665	4.132
15–24, rural, superior → Familiar no remunerado	0.015	0.654	4.064
Mujer, 15–24, urbana, superior, sin NIT → Familiar no remunerado	0.016	0.653	4.058
Mujer, 15–24, urbana, superior → Familiar no remunerado	0.018	0.633	3.929
15–24, rural, sin NIT → Familiar no remunerado	0.018	0.621	3.859
15–24, rural → Familiar no remunerado	0.018	0.601	3.732

Las reglas revelan que mujeres jóvenes (15–24 años), especialmente en áreas rurales y sin NIT, tienen alta probabilidad de ser trabajadoras no remuneradas.

Estas combinaciones presentan lift > 4 y alta confianza, evidenciando un perfil estructural de informalidad.

RESULTADOS

DETERMINANTES DE TENER O NO TENER NIT EN EL LUGAR DE TRABAJO

Regla	Support	Confidence	Lift
Hombre, 40–59, otros, asalariado → NIT: Sí	0.016	0.833	4.064
Hombre, 40–59, urbano, otros, asalariado → NIT: Sí	0.016	0.831	4.055
40–59, otros, asalariado → NIT: Sí	0.026	0.824	4.019
40–59, urbano, otros, asalariado → NIT: Sí	0.026	0.822	4.009
Mujer, 40–59, otros, asalariado → NIT: Sí	0.010	0.809	3.947
Mujer, 40–59, urbano, otros, asalariado → NIT: Sí	0.010	0.807	3.939
Mujer, 25–39, urbano, otros, asalariado → NIT: Sí	0.032	0.767	3.744
Mujer, 25–39, otros, asalariado → NIT: Sí	0.032	0.767	3.743
25–39, urbano, otros, asalariado → NIT: Sí	0.067	0.757	3.696
25–39, otros, asalariado → NIT: Sí	0.068	0.751	3.665

Esta regla en específico es destacable, dado que muestra que mujeres mayores con educación no formal logran obtener formalización.

- Hombres de 40–59 años, con formación técnica "no tradicional" y ocupación asalariada, presentan alta probabilidad de tener NIT (confianza > 80%, lift \approx 4). La experiencia práctica pesa más que la educación formal.
- Mujeres y jóvenes (25–39) tienen menor probabilidad de formalización, pero algunos perfiles técnicos urbanos alcanzan > 76% de confianza, mostrando potencial para políticas focalizadas.

CONCLUSIONES

- Informalidad alta en mujeres jóvenes rurales, incluso con educación superior (hasta 89% sin remuneración).
- Formalización destacada en hombres de 40-59 años con formación técnica no tradicional (confianza > 83% con NIT).
- Brechas persistentes: las mujeres rurales siguen siendo el grupo más vulnerable en el acceso a empleo formal.
- Las reglas de asociación permitieron detectar perfiles complejos que no surgen de análisis convencionales.

RECOMENDACIONES

- **Mujeres en áreas rurales:** Se recomienda implementar programas de primer empleo con incentivos fiscales que motiven a los empleadores a contratar jóvenes, especialmente mujeres. Esto debe ir acompañado de capacitación técnica orientada a las necesidades productivas locales
- **Reconocimiento a trabajadores con experiencia:** Debe existir mayor valoración a los saberes prácticos adquiridos fuera de un sistema educativo formal.
- **Enfoque segmentado y territorial:** Los resultados muestran que la informalidad no tiene un efecto en todos por igual. Por ello, se recomienda abandonar soluciones genéricas y avanzar hacia políticas públicas diferenciadas, focalizadas en los perfiles vulnerables detectados, y ajustadas al contexto regional.



**MUCHAS
GRACIAS**