

Taller 01 - Estimación del Salario de los Ocupados en el Area Metropolitana del Valle de Aburrá

Estudiante

Jesús Andrés Álvarez Alvarado

Docente

Osmar Leandro Loaiza Quintero

Asignatura

Econometría Avanzada



UNIVERSIDAD
NACIONAL
DE COLOMBIA

Sede Medellín
30 de marzo de 2024

Tabla de Contenidos

1	Cruce, Limpieza de Datos y Elección de Variables	3
1.1	Variables	3

Lista de Figuras

Lista de Tablas

1 Cruce, Limpieza de Datos y Elección de Variables

Lo primero será seleccionar las variables, filtrar por área y en general, alistar los datos para ingresarlos al modelo.

1.1 Variables

Las variables seleccionadas se encuentran en dos módulos, *Características generales, seguridad social en salud y educación* y *Ocupados*. La descripción completa de los módulos y en detalle de cada variable se encuentra en los metadatos (diccionario). Las variables son:

Del módulo de *Características generales, seguridad social en salud y educación*:

- DIRECTORIO, SECUENCIA_P, HOGAR, ORDEN: son los identificadores de la vivienda, hogar y personas. Estas variables se utilizan para unir las variables de cada módulo
- AREA: es la variable que permite filtrar los datos por área. En este caso, como se trabaja con el Área Metropolitana del Valle de Aburrá, se filtrará por el valor 05 de la variable
- P3271: ¿cuál es el sexo al nacer? Para capturar si hay brechas salariales asociadas al sexo
- P6040: ¿cuántos años cumplidos tiene? Para capturar diferencias salariales entre edades
- P6050: ¿cuál es el parentesco con el jefe del hogar? Para capturar cómo los roles en el hogar pueden influenciar el salario
- P6080: ¿se reconoce dentro de alguna etnia? Para capturar brechas salariales entre diferentes etnias
- P6070: estado civil
- P3042: ¿cuál es el mayor nivel educativo alcanzado y el último grado o semestre aprobado?

Del módulo de *Ocupados*:

- DIRECTORIO, HOGAR, SECUENCIA_P, ORDEN: son los identificadores de la vivienda, hogar y personas. Estas variables se utilizan para unir las variables de cada módulo
- AREA: es la variable que permite filtrar los datos por área. En este caso, como se trabaja con el Área Metropolitana del Valle de Aburrá, se filtrará por el valor 05 de la variable

- INGLABO: será la variable respuesta, ya que hace referencia a los ingresos laborales mensuales de la persona
- P6440: ¿para realizar este trabajo tiene usted algún tipo de contrato? Pretende ser una aproximación para capturar información de formalidad e informalidad
- P6920: ¿está cotizando actualmente a un fondo de pensiones? También intenta aproximar a las personas que son formales e informales. Pues, es poco común que los informales coticen a pensión. No se utiliza afiliación a salud, pues, existe una gran cobertura de salud bajo regimen subsidiado en Colombia, lo que no permite segmentar correctamente formales de informales
- P6960: ¿cuántos años lleva afiliado al fondo de pensiones? Dado que no se pregunta directamente cuántos años de experiencia tiene la persona, se intenta aproximar la experiencia por medio de los años cotizados a pensión. Esto claramente excluirá a la población informal, pero es la variable más próxima a la experiencia