Parte I

Una vez unidas las bases de datos de las encuestas individual y de hogar correspondiente al primer trimestre de los años 2004 y 2024 de acuerdo al código de identificación del hogar y al número de hogar, primero revisamos que el rango de valores que tomaran las variables tuviese sentido. Para ello realizamos algunas modificaciones en la base de datos de manera que las variables pertenecieran a un tipo de formato específico, por ejemplo y sin pérdida de generalidad, para la variable Edad (CH06) los valores eran números enteros a excepción de los menores a un año. Para estas observaciones la variable indicaba en formato string “Menor a 1 año”. Dado que el trabajo se enfocaba en predecir y analizar la desocupación, decidimos deshacernos de esas observaciones: eran pocas y no contribuían al análisis de la temática de interés.

Para el caso de los ingresos, tanto del Ingreso Familiar Total y del Ingreso Per Cápita Familiar (ITF e IPCF respectivamente y a partir de ahora) decidimos filtrar la base para los dos diferentes años, ya que eliminar outliers de una base que contiene a la vez datos sobre los ingresos corrientes de agentes y hogares para dos períodos alejados en el tiempo no tendría sentido debido al efecto inflacionario.

Gráfico 1: ITF para los años 2004 y 2024 en pesos corrientes de cada año ***con outliers***

Gráfico, Histograma

Descripción generada automáticamente

Para el año 2004 contabámos con 2547 observaciones iniciales y para el año 2024 con 2098 observaciones. Decidimos eliminar de la muestra los reportes que indicaban que el hogar tenía ingresos nulos. Para el año 2004, menos del 1% de los hogares reportó ingresos nulos, sin embargo, para el año 2024 estos reportes representaron el 16% de la muestra, un porcentaje muy alto. Sin embargo nos resultó un sinsentido mantener estos reportes en la muestra debido a la imposibilidad de vivir con ingresos nulos.

Gráfico 2: ITF para los años 2004 y 2024 en pesos corrientes de cada año ***sin outliers***

Gráfico, Histograma

Descripción generada automáticamente