

## Parte 1:

- 1) Se considera “persona pobre” a aquellas que están por debajo de la línea de pobreza. Estar por debajo de la línea de pobreza implica no satisfacer la canasta básica total, que incluye la canasta básica alimentaria más los bienes y servicios no alimentarios.
- 2) En el inciso a) filtramos a la región 1 GBA. Luego en el inciso b) corregimos cualquier información que podría contener datos irracionales. Por eso filtramos edad, ingresos y horas trabajadas a número no negativos, pero tampoco irracionalmente altos. En el c), primero creamos las dos variables “varón” y “mujer” y creamos el gráfico de barra. En el d) armamos la matriz de correlación con la primer opción que mostraba el link de la consigna. Para ello utilizamos parte de la biblioteca de Seaborn. Para el e) primero identificamos de la variable Estado la cantidad de ocupados desocupados e inactivos, filtramos estos 3 (dejando fuera categoría 4: menor de 10 años) y luego calculamos su media de ingresos. Por último para la f, utilizamos el excel adjunto creamos la columna adulto\_equiv que tiene los valores de dieta para su sexo y edad y luego con groupby los sumamos esta columna para las personas que pertenecían al mismo hogar (coincidía num\_hogar) y esto quedo en ad\_equiv\_hogar.Avc
- 3) Le dimos el nombre de ITF al ingreso total familiar. Luego pedimos que nos filtre a ITF entre valores iguales a cero y valores distintos a cero y que guarde los datos en dos archivos distintos llamados respondieron y no\_respondieron. Por último pedimos que cuenta y haga un print de las personas que no respondieron.
- 4) Hacemos el producto entre la canasta básica y los adultos equivalentes por hogar para encontrar el valor de ingresos que cada hogar necesita para no ser pobre.
- 5) Creamos un condicional que tome valor 1 si  $ITF < \text{ingreso necesario}$  y valor cero en caso contrario. A esta variable la llamamos pobre y es justamente la que indica si un hogar está por debajo de la línea de pobreza.

## Parte 2:

- 1) Definimos y agrupamos todas las variables a eliminar en una sola función. Luego con drop columns las eliminamos.
- 2) Agregamos intercepto (columna de unos). Luego definimos a pobre como variable dependiente y a todas las variables excepto a pobre como independientes. Pedimos test size = 0,3 y utilizamos la semilla pedida en la consigna.