**Análisis de Correspondencia Simple (ACS)**

Con el objetivo de determinar si existe alguna asociación entre los grupos resultantes del análisis Cluster y las etiquetas con las que fueron construidos los individuos representativos de la base de datos, se implementó la técnica de correspondencias simples.

El presente análisis sólo tiene sentido si existe alguna dependencia entre las variables utilizadas en la tabla de contingencia. Para testear lo anterior, se utilizó el test Chi-cuadrado de Pearson. Para todos los casos (los grupos de cluster contra cada una de las etiquetas de los individuos representativos) se rechazó la hipótesis nula de independencia, lo cual habilita la implementación de un ACS.

Es importante destacar que, a priori, es posible establecer que los pesos para las modalidades localidad, sexo, nivel educativo, y edad van a ser los mismos (es decir, el valor que nos devuelve R como colmass). Esto es así puesto que, por construcción de la base, se tiene la misma cantidad de individuos en cada modalidad, por lo que, los pesos van a ser los mismos en cada una de ellas. No se hará referencia a este hecho nuevamente en lo que resta del análisis.

Otro punto importante a subrayar es el hecho de que, también a priori, se sabe que los grupos en Cluster son de diferentes tamaños (en cantidad de individuos), siendo el grupo 1 y 4 particularmente altos. Esto significa que, los grupos que van a tener mayor peso van a ser siempre primero el 1 y luego el 4. Tampoco se volverá a mencionar este hallazgo en lo que sigue.

**Grupos – Localidad y Sexo**

En una primera instancia, se implementó ACS con los grupos de cluster y una variable cualitativa cuyas modalidades son en función del sexo y la localidad donde reside el individuo. De esta manera, se tienen 4 modalidades: interior-hombre (IH), interior- mujer (IM), Montevideo-hombre (MH), Montevideo-mujer (MM).

La tabla de contingencia se presenta a continuación:

**Tabla 4.1.1**



**Tabla 4.1.2**



La tabla 4.1.2 muestra que con 2 dimensiones se logra captar casi el 100% de la inercia de la nube de puntos proyectada, lo cual quiere decir que reducir dimensiones no implica casi pérdida alguna de información.

**Tabla 4.1.3**



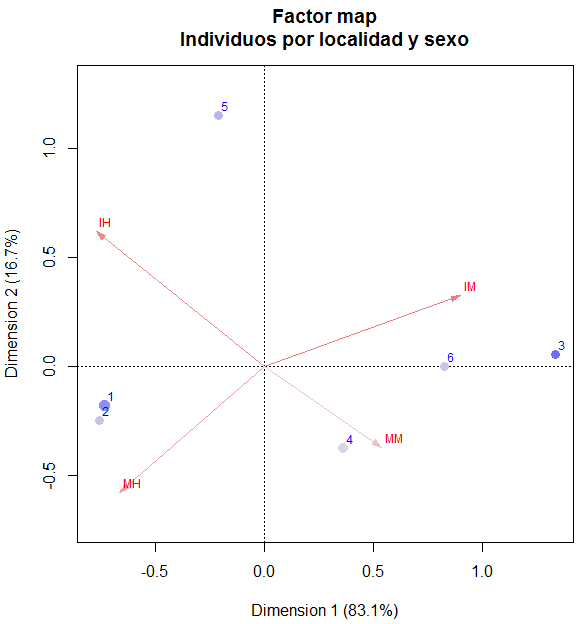
En lo que respecta a contribución a la inercia, la misma está dada primordialmente por el grupo 3. La distancia chi-cuadrado es la distancia de cada grupo al centroide. cuando la misma es cercana a uno, la modalidad está mejor representada. Este último punto es de suma importancia ya que para interpretar los puntos, los mismos deben estar bien proyectados. El grupo 3 es totalmente dominante en lo que respecta a calidad de representación en el factor map.

**Tabla 4.1.4**



Para las modalidades de localidad-sexo, si bien no existen grande diferencias, vemos que Interior-Hombre e Interior-Mujer son las que están mejor representadas, al mismo tiempo que son las que más contribuyen a la inercia.

A continuación se presenta el factor map asociado:



No parece haber ningún punto que esté especialmente alejado de la nube de puntos, por lo cuál no es preciso quitar ningún punto e incluirlo luego como suplementario.

El factor map sugiere que, mientras que el primer eje separa por localidad, el segundo eje separa por sexo.

**Grupos – Nivel educativo**

En una segunda instancia nos proponemos investigar si existe alguna asociación entre los grupos formados y el nivel educativo de los individuos. A continuación se presentan la tabla de contingencia y la tabla de valores propios y varianzas:

**Tabla 4.2.1**



**Tabla 4.2.2**



La tabla 4.2.2 muestra que la primera dimensión capta el 83% dela varianza de los datos proyectados mientras que la segunda capta el 17%. Entre las dos, se acumula el 100% de la varianza.

**Tabla 4.2.3**



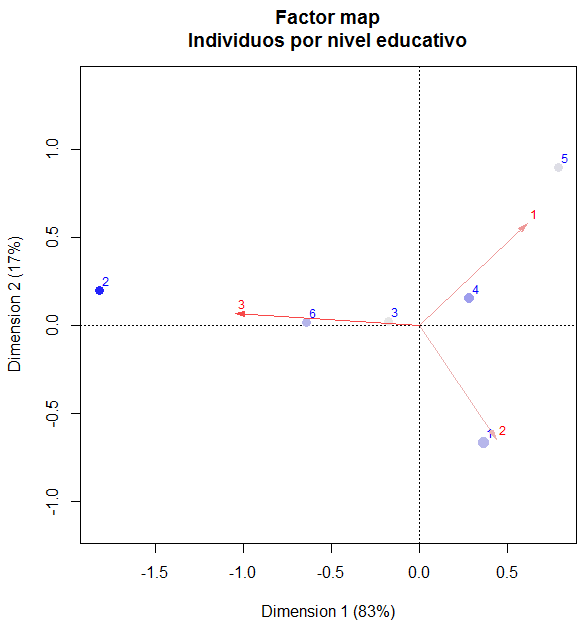
En esta oportunidad, la modalidad que está mejor representada es el grupo 2, el cuál es totalmente dominante, ya que es el que toma el valor más alto en la distancia chi-cuadrado al centroide. En lo que respecta a inercia, el grupo 2 es también el que más aporta.

**Tabla 4.2.4**



Para las modalidades de la variable nivel educativo, es nivel terciario la que queda mejor representada a la vez que la que contribuye más a la inercia.

El factor map se presenta a continuación:



A partir del análisis visual se ve que el nivel educativo terciario es el que más contribuye a la conformación del eje 1 mientras que, para el eje2, no resulta tan claro si es la modalidad 1 o 2 de la variable nivel educativo la que mas contribuye a su conformación.

En esta oportunidad, el factor map sugiere que hay una clara asociación entre el grupo 1 y el nivel educativo 2 (secundaria). A la vez que los grupos 3, 6, y 2 se asocian más claramente con el nivel educativo 3 (terciario). Es importante destacar la concordancia de lo que revela el factor map con la caracterización de los grupos hecha en cluster ***(No se si podemos interpetar esto pero parecería que tiene sentido……………….).***

**Grupos – Edad**

Por último, se implementó correspondencia simple con los grupos de cluster y los tramos de edad. Las tablas de contingencia y de varianzas son presentadas a continuación:

**Tabla 4.3.1**



**Tabla 4.3.2**



La dimensión 1 capta el 82% de la varianza de los datos proyectados mientras que la dimensión 2 el 18%. Entre las dos, se acumula el total de 100%.

**Tabla 4.3.3**



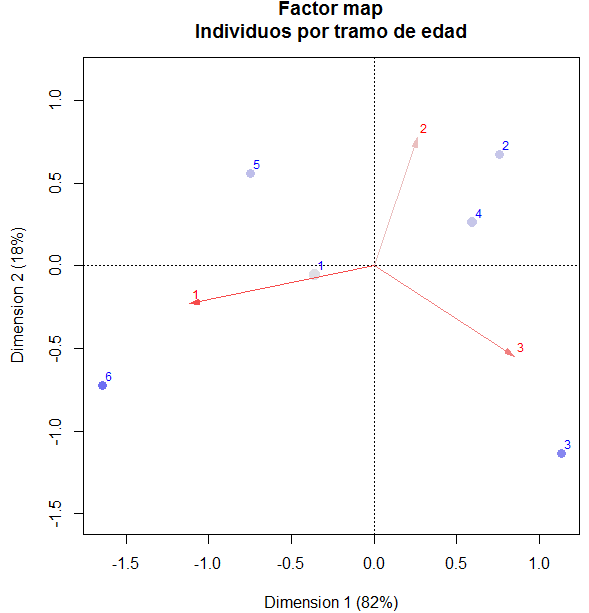
La tabla 4.3.3 evidencia que los grupos 6 y 3 son los que están mejor representados puesto que son los más próximos al círculo unidad. Asimismo, son los que contribuyen más a la inercia.

**Tabla 4.3.4**



Por su parte, para los tramos de edad, el tramo de edad 1 es la modalidad que queda mejor representada a la vez que es la que más contribuye a la inercia.

El factor map se presenta a continuación:



Atendiendo al análisis visual del factor map, la modalidad 1 de la variable edad (20-29 años) es la que más contribuye a la conformación del eje 1 y la modalidad 2 (30-49 años) es la que más lo hace a la conformación del eje 2. El efecto de la modalidad 3 en los ejes no parece demasiado claro.

**Conclusiones**

* Si bien el grupo 4 fue caracterizado en Cluster como un grupo de observaciones heterogéneas, el presente análisis sugiere que está asociado a la modalidad Montevideo-Mujeres.
* El grupo 1, que había sido caracterizado como de observaciones heterogéneas, parecería guardar una relación con la modalidad nivel educativo 2, mientras que, los grupos 3,6, y 2 revelan la información extraída en análisis Cluster de que están asociados a la modalidad nivel educativo terciario.
* En lo que respecta a las edades, no es posible exponer ninguna conclusión clara en lo que respecta a asociación con los grupos.
* En definitiva, los grupos parecen diferenciarse más que nada en función del nivel educativo.

Combinando los resultados encontrados en la aplicación de las dos metodologías (componentes principales y correspondencias simples), puede esbozarse una importante conclusión sobre la estructura del mercado laboral uruguayo, la cual hace hace referencia a la movilidad que proporciona la educación en el ámbito laboral.

Parecería ser que la calidad de la inserción laboral de los individuos está estrechamente relacionada en primera instancia con el nivel educativo alcanzado en sus vidas y, en menor medida, su edad (entre los individuos con niveles educativos altos, son los jóvenes los que parecen tener una inserción laboral un poco más endeble).