



1. Solución:

$$\text{Si } L = \begin{bmatrix} .9 & .7 & .5 \end{bmatrix}' \text{ y } \Psi = \begin{bmatrix} .19 & 0 & 0 \\ 0 & .51 & 0 \\ 0 & 0 & .75 \end{bmatrix}$$

$$\text{Entonces: } \rho = LL' + \Psi = \begin{bmatrix} 0.81 & 0.63 & 0.45 \\ 0.63 & 0.49 & 0.35 \\ 0.45 & 0.35 & 0.25 \end{bmatrix} + \begin{bmatrix} .19 & 0 & 0 \\ 0 & .51 & 0 \\ 0 & 0 & .75 \end{bmatrix} = \begin{bmatrix} 1 & 0.63 & 0.45 \\ 0.63 & 1 & 0.35 \\ 0.45 & 0.35 & 1 \end{bmatrix}$$

La comunalidad:  $h = [0.81 \ 0.49 \ 0.25]$ , la primer y segunda variable son las que más contribuyen a la varianza para el factor.

La primera variable es la que lleva el mayor peso en las cargas, y por lo tanto en la interpretación del factor común, esto también lo podemos observar en la matriz de covarianzas donde la variable 1 con la 2 y 3 son con las que mayor presenta correlación.

$$\text{corr}(Z_1, F1) = \text{cov}(Z_1, F1) = l_{1,1} = 0.9$$

$$\text{corr}(Z_2, F1) = \text{cov}(Z_2, F1) = l_{2,1} = 0.7$$

$$\text{corr}(Z_3, F1) = \text{cov}(Z_3, F1) = l_{3,1} = 0.5$$



2. Solución:

Los valores y vectores propios de  $\rho$  son:  $[1.96 \ 0.67 \ 0.35]$ ,  $\begin{bmatrix} -0.62 & .21 & 0.74 \\ -0.59 & .49 & -0.63 \\ -0.5 & -0.84 & -0.17 \end{bmatrix}$

La proporción de varianza explicada por 1 factor es 0.65.

Las cargas son:  $L = [-0.87 \ -0.83 \ -0.71]'$ ,  $LL' = \begin{bmatrix} 0.76 & 0.72 & 0.62 \\ 0.72 & 0.69 & 0.35 \\ 0.62 & 0.59 & 0.5 \end{bmatrix}$

La comunalidad:  $h = [0.76 \ 0.69 \ 0.5]$

Y, por lo tanto:

$$\Psi = \rho - LL' = \begin{bmatrix} 1 & 0.63 & 0.45 \\ 0.63 & 1 & 0.35 \\ 0.45 & 0.35 & 1 \end{bmatrix} - \begin{bmatrix} 0.76 & 0.72 & 0.62 \\ 0.72 & 0.69 & 0.35 \\ 0.62 & 0.59 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.23 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.49 \end{bmatrix}, \text{ entonces:}$$

$$\hat{\rho} = LL' + \Psi = \begin{bmatrix} 1 & 0.72 & 0.62 \\ 0.72 & 1 & 0.59 \\ 0.62 & 0.59 & 1 \end{bmatrix}$$

Como puede observar, las cargas son mas grandes y con signos diferentes, las comunalidades también son más grandes en magnitud y por lo tanto las varianzas especificas son menores y la diferencia entre la matriz de covarianzas original y la estimada, es pequeña pero no exactamente igual a 0, es decir son ligeramente diferentes.

$$\rho - \hat{\rho} = \begin{bmatrix} 0 & 0.09 & 0.17 \\ 0.09 & 0 & 0.24 \\ 0.17 & 0.24 & 0 \end{bmatrix}$$



3. Solución:

Si  $\Sigma = LL' + \Psi$ , entonces:

$$1 = l_{3,1} + \psi_3, \quad .4 = l_{1,1}l_{2,1}, \quad .9 = l_{1,1}l_{3,1}, \quad .7 = l_{2,1}l_{3,1}$$

$$\text{De lo anterior: } l_{2,1} = \frac{.4}{.9}l_{3,1}, \quad l_{3,1} \left( \frac{.4}{.9}l_{3,1} \right) = .7 \rightarrow l_{3,1} = \pm 1.255$$

Entonces:  $\psi_3 = 1 - l_{3,1} = -0.255$ , como  $\psi_3 < 0$ , por lo tanto, no es admisible.



## 4. Solución:

Para las 50 observaciones para las 7 variables (crecimiento de ventas, rentabilidad de ventas, ventas de nuevas cuentas, creatividad, razonamiento mecánico, razonamiento abstracto y capacidad matemática), se aplicó la prueba de esfericidad de Barlett y se rechazó la hipótesis nula por lo cual se aplicó el método de factores, después mediante un modelo ortogonal de factores, la solución por máxima verosimilitud con una rotación varimax es:

Para  $m=2$  factores

$$\text{Las cargas: } L = \begin{bmatrix} .85 & .45 \\ .86 & .41 \\ .71 & .6 \\ .14 & .98 \\ .5 & .52 \\ .61 & .05 \\ .94 & .27 \end{bmatrix}$$

y la diagonal de la varianza específica  $\psi = [.06 \ .07 \ .12 \ .00 \ .47 \ .61 \ .02]$

El primer factor parece englobar las variables 7,2 y 1 que están relacionadas con una habilidad matemática para las ventas, entonces será el factor de habilidad matemática. El segundo factor es de la variable 4 y le podemos llamar factor de creatividad.

$$R = LL' + \Psi \rightarrow R - (LL' + \Psi) = \begin{bmatrix} 0 & 0 & 0 & 0 & .04 & .12 & 0 \\ 0 & 0 & -.03 & 0 & .09 & -.09 & 0 \\ 0 & -.03 & 0 & 0 & -.03 & .16 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .04 & .09 & -.03 & 0 & 0 & .04 & -.04 \\ .12 & -.09 & .16 & 0 & .04 & 0 & -.03 \\ 0 & 0 & 0 & 0 & -.04 & -.03 & 0 \end{bmatrix}$$

Se realizó también una prueba para probar  $H_0: R = LL' + \Psi$ , con  $m=2$  factores, la cual se rechaza ( $p\text{-valor}=8e-22$ ) y la varianza acumulada en este caso es de 0.8 (Bajo este criterio, podría ser suficiente).

Para  $m=3$  factores

Las cargas:  $L = \begin{bmatrix} .79 & .37 & .43 \\ .91 & .31 & .18 \\ .65 & .54 & .43 \\ .25 & .96 & .01 \\ .54 & .46 & .20 \\ .29 & .05 & .95 \\ .91 & .17 & .29 \end{bmatrix}$

y la diagonal de la varianza especifica  $\psi = [.03 \ .03 \ .08 \ .00 \ .44 \ .00 \ .03]$

El primer factor nuevamente engloba las variables 7,2 y 1 que están relacionadas con una habilidad matemática para las ventas, entonces será el factor de habilidad matemática. El segundo factor es de la variable 4 y le podemos llamar factor de creatividad y el tercer factor la variable 6, el factor de la abstracción.

$$R = LL' + \Psi \rightarrow R - (LL' + \Psi) = \begin{bmatrix} 0 & 0 & -.02 & 0 & .01 & 0 & 0 \\ 0 & 0 & 0 & 0 & .06 & 0 & 0 \\ -.02 & 0 & 0 & 0 & -.05 & 0 & .02 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .01 & .06 & -.05 & 0 & 0 & 0 & -.06 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .02 & 0 & -.06 & 0 & 0 \end{bmatrix}$$

Puede observar que si la compara con la aproximación con  $m=2$  factores, esta aproximación es ligeramente mejor.

Se realizó también una prueba para probar  $H_0: R = LL' + \Psi$ , con  $m=3$  factores, la cual se rechaza ( $p\text{-valor}=1e-13$ ) y la varianza acumulada en este caso es de 0.9. es decir podemos ver que aunque rechaza, mejora ligeramente, pero no es suficiente, sin embargo con el criterio de la varianza acumulada que es aproximadamente el 90% y que la aproximación fue bastante buena, se toma  $m=3$  factores y se calculan los factor scores mediante el método de mínimos cuadrados ponderados, los cuales se pueden observar en la Figura 4.1, note que no hay algún patrón visible en ninguna de las proyecciones, ya que no hay algún conjunto de puntos que claramente se separe del resto, al contrario todos forman una nube de puntos.

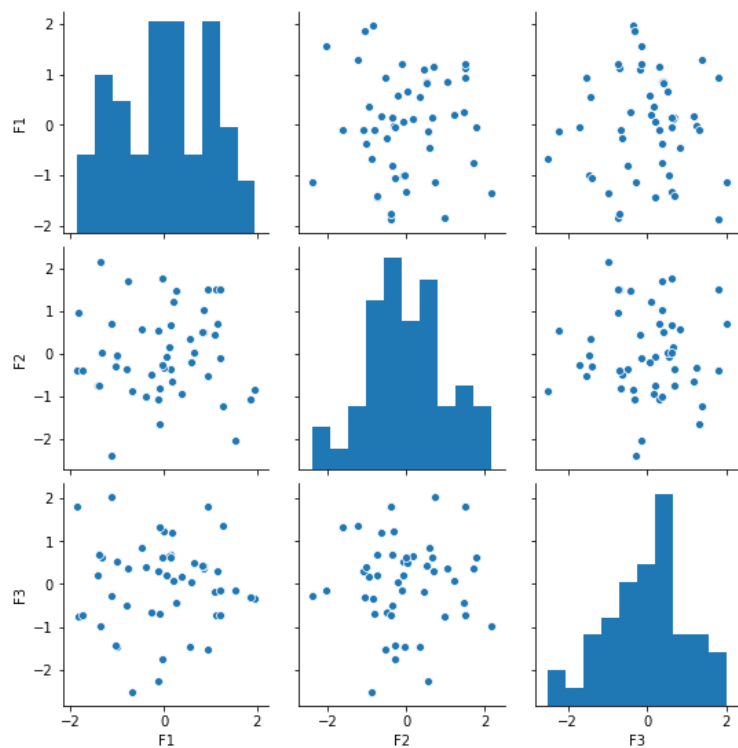


Figura 4.1: matriz de dispersión de los factor scores.



## 5. Solución:

Para el subconjunto de datos del proyecto SAPA, los cuales estaban completos ( $n=2236$ ) para las 25 variables (A1...O5) se utilizó el análisis de factores para agrupar elementos de naturaleza similar (previo se realizó la prueba de esfericidad de Barlett la cual se rechazó ( $p\text{-valor}=0$ ) y por lo tanto se aplicó el análisis de factores), el número de factores ( $m=5$ ) se determinó mediante el método del screenPlot (Figura 5.1), con este numero de factores se tiene una varianza acumulada del 53%.

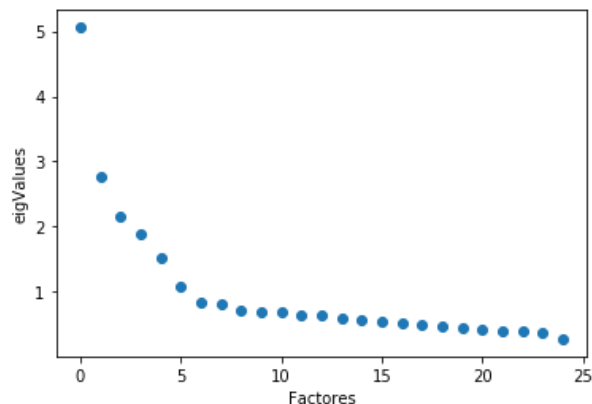


Figura 5.1: Screenplot, punto de cambio en 5, es donde hay un cambio de pendientes.

Se determinaron las cargas (Figura 5.2) y se formaron los 5 factores, los cuales, el primero (F1) esta relacionado a la irritabilidad y enojo, el factor 2 (F2) esta relacionado a la capacidad para relacionarse, el tercer factor (F3) está relacionado con el trabajo y como se desempeña, el cuarto (F4) esta relacionado a la empatía hacia otros, el quinto (F5) se relaciona con abundar en temas y de reflexión.

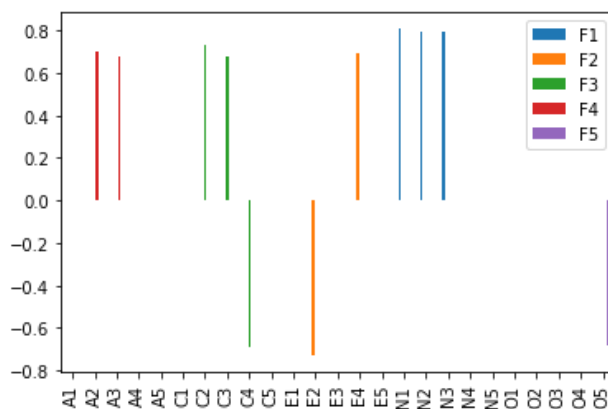


Figura 5.2: Cargas (L) de los factores y a que preguntas corresponden.

Posteriormente se obtuvieron los factor scores y mediante gráficos de dispersión (Figura 5.3) se buscaron datos atípicos, en la mayoría se observa un ligero sesgo hacia la izquierda así que ese fue un punto para buscar los datos atípicos, es decir los que mas se alejan o se separan del resto.

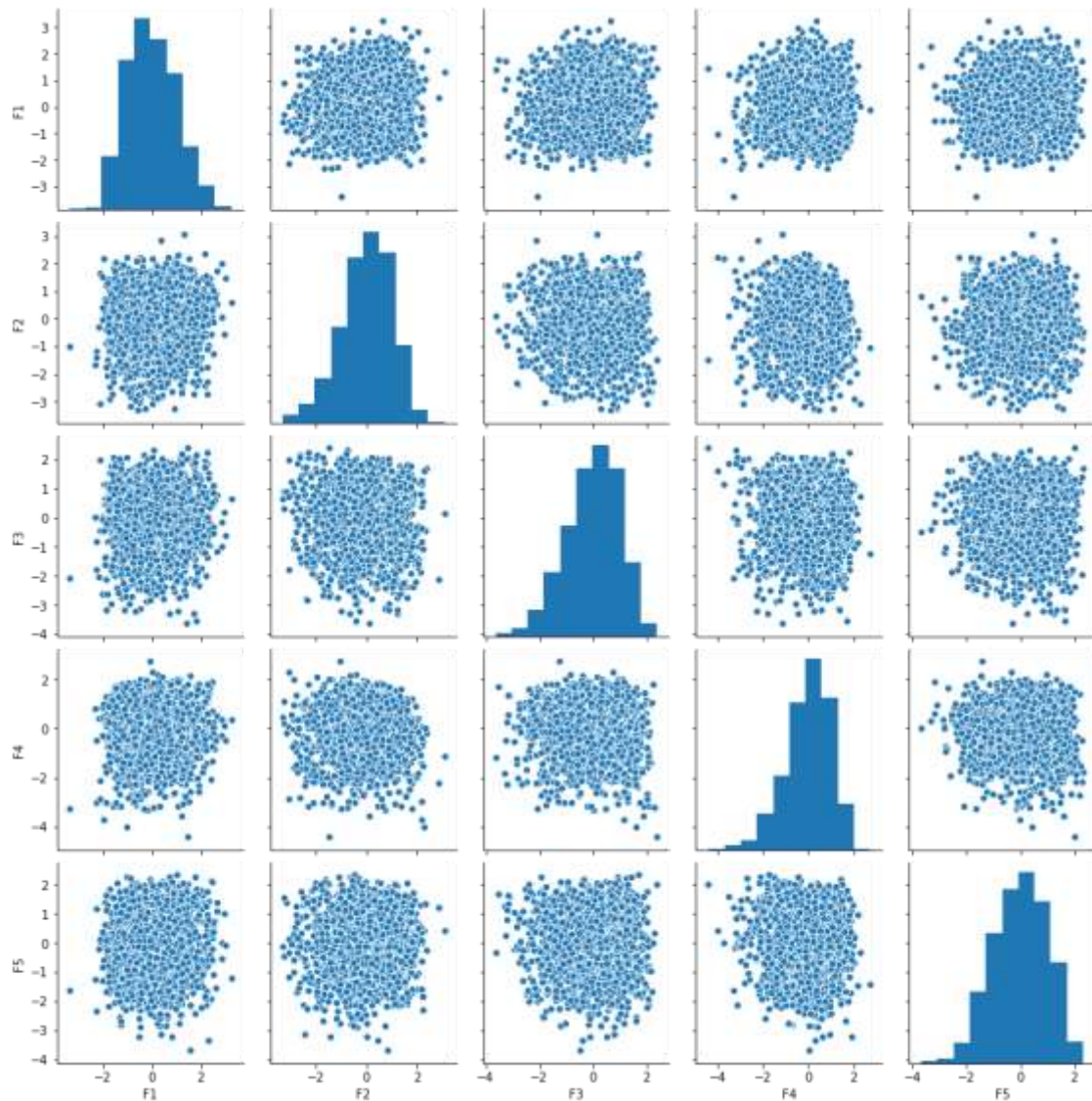


Figura 5.3: Matriz de dispersión de los factor scores. En esta figura podemos observar como se distribuye cada factor y su dispersión en diferentes proyecciones para encontrar patrones o datos atípicos.



Se encontró principalmente 4 datos atípicos relacionados a los factores 1 y 4 (Figura 5.4), estas observaciones son de casos extremos, es decir son sujetos que respondieron a todo o en su mayoría o 1 o 6 (extremos).

En el caso de la observación 525, observe el contraste en las respuestas, para el cuarto factor básicamente responde que es muy indiferente a los sentimientos de los demás y no sabe cómo tratar con ellos, para el tercer factor respondió que le gusta hacer su trabajo y hacerlo muy bien sin perder el tiempo. Para el segundo factor no habla mucho y tiene dificultad para interactuar con los demás, para primer factor tiene cambios de humor extremos, se enoja e irrita fácilmente, y respecto al quinto factor es una persona de muchas ideas que le gusta leer material complejo y llevar la conversación a ese nivel, es decir le gusta profundizar en el tema.

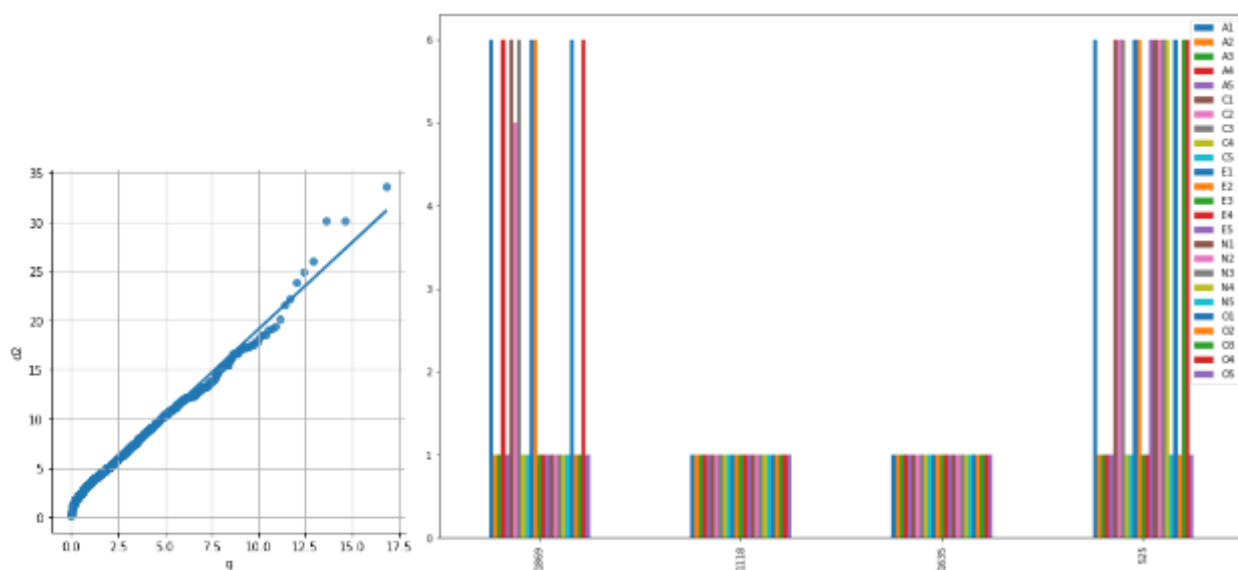


Figura 5.4: a) Gráfico chi-cuadrado, b) Datos atípicos, 4 observaciones que más se alejan de la recta azul (1869, 1118, 1635, 525).