

INFORME DATOS “ENTREVISTA”

Navarro Castillo, Alejandra
Análisis de Datos
01 abril 2021

Introducción y objetivo

En una entrevista de trabajo se han recogido datos de las calificaciones obtenidas por los solicitantes del puesto de trabajo en diferentes ámbitos. Se han medido 15 variables con una escala números enteros del 0 al 10.

El resumen de los datos lo podemos observar en la siguiente tabla. (La variable *Obs* es la etiqueta de cada candidato al puesto de trabajo, es decir, el número de observación)

	mean	sd	IQR	0%	25%	50%	75%	100%	n
Obs	24.500000	14.000000	23.50	1	12.75	24.5	36.25	48	48
Form.of.letter.of.application	6.000000	2.673749	5.00	0	4.00	6.0	9.00	10	48
Appearance	7.083333	1.966023	2.25	3	6.00	7.0	8.25	10	48
Academic.ability	7.083333	1.987550	2.00	2	6.00	7.0	8.00	10	48
Likeability	6.145833	2.805690	4.00	0	4.00	7.0	8.00	10	48
Self.confidence	6.937500	2.418072	4.00	1	5.00	8.0	9.00	10	48
Lucidity	6.312500	3.170048	5.00	0	4.00	8.0	9.00	10	48
Honesty	8.041667	2.534514	2.00	0	8.00	9.0	10.00	10	48
Salesmanship	4.854167	3.439381	6.00	0	2.00	4.5	8.00	10	48
Experience	4.229167	3.308529	5.25	0	2.00	3.0	7.25	10	48
Drive	5.312500	2.947457	5.00	0	3.00	5.0	8.00	10	48
Ambition	5.979167	2.935401	5.25	0	3.75	6.0	9.00	10	48
Grasp	6.250000	3.035254	4.00	0	4.00	7.0	8.00	10	48
Potential	5.687500	3.183443	4.25	0	3.75	6.0	8.00	10	48
Keeness.to.join	5.562500	2.657036	3.25	0	4.00	5.0	7.25	10	48
Suitability	5.958333	3.300279	6.25	0	3.00	6.0	9.25	10	48

Observamos a través del resumen de los datos cómo son las variables que estamos estudiando. La mayoría de ellas tienen su media alrededor del 5, pero destaca la variable *Honesty* cuya media es mucho más alta que la de las demás variables.

En cuanto a la dispersión de las variables, las variables con mayor desviación típica son *Suitability*, *Salesmanship*, *Experience*, lo que indica mayor diferencia de capacidad de los candidatos en esos ámbitos. Las características en las cuales los candidatos tienen un perfil más parecido son *Appearance* y *Academic.ability*.

Dado que se ha medido un gran número de variables, queremos reducir este número para poder entender e interpretar mejor los datos obtenidos. Por ello el objetivo de este informe es observar e identificar las características comunes que se encuentran entre todas las variables que se midieron en el estudio inicial.

Método

El método que hemos empleado para realizar el análisis de los datos va a ser el Análisis Factorial que trata de reducir el número de variables a partir de grupos homogéneos de variables. Los grupos se forman con las variables que más correlación tienen entre sí y procuran que cada grupo sea independiente a los demás.

Para empezar, hemos analizado la matriz de correlaciones entre las variables para verificar si es viable hacer un análisis factorial.

Matriz de correlación de las 6 primeras variables.

	Form.of.letter.of.application	Appearance	Academic.ability	Likeability	Self.confidence	Lucidity
Form.of.letter.of.application	1.000	0.239	0.044	0.306	0.092	0.228
Appearance	0.239	1.000	0.123	0.380	0.431	0.371
Academic.ability	0.044	0.123	1.000	0.002	0.001	0.077
Likeability	0.306	0.380	0.002	1.000	0.302	0.483
Self.confidence	0.092	0.431	0.001	0.302	1.000	0.808
Lucidity	0.228	0.371	0.077	0.483	0.808	1.000

Como se aprecia en la tabla, hay variables altamente correladas con otras, así que hemos procedido a hacer en análisis factorial.

Hemos elegido en este análisis 4 factores, lo que corresponde a representar un 74,4% de la variabilidad total de los datos. Estos factores que hemos extraído y que son las características comunes se pueden definir como sigue:

- *Factor1* representa “Motivación y habilidades sociales” de los candidatos.
- *Factor2* representa “Idoneidad y preparación” de los candidatos para el puesto de trabajo concreto.
- *Factor3* representa “Personalidad” de los candidatos.
- *Factor4* representa “Titulaciones académicas”

Así hemos conseguido resumir las 15 variables principales en estos cuatro últimos factores.

Apéndice

Código del Análisis Factorial en RStudio.

```
3 # Cargar los datos
4 load("~/Desktop/AnalisisdeDatos/2.Analisis_factorial/Entrevista.RData")
5
6 # Describir las variables
7 names(Entrevista)
8 str(Entrevista)
9
10 summary(Entrevista)
11 numSummary(Entrevista)
12
13 # Verificar que es viable hacer un análisis factorial
14 Mcor <- round(cor(Entrevista[2:16]),3)
15 View(Mcor)
16 Mcor[1:6,1:6]
17
18 # Extraer los factores
19 eigen(Mcor)
20
21 AnalisisF <- factanal(Entrevista[-1], factors=3, rotation="varimax", scores="none")
22 AnalisisF
```