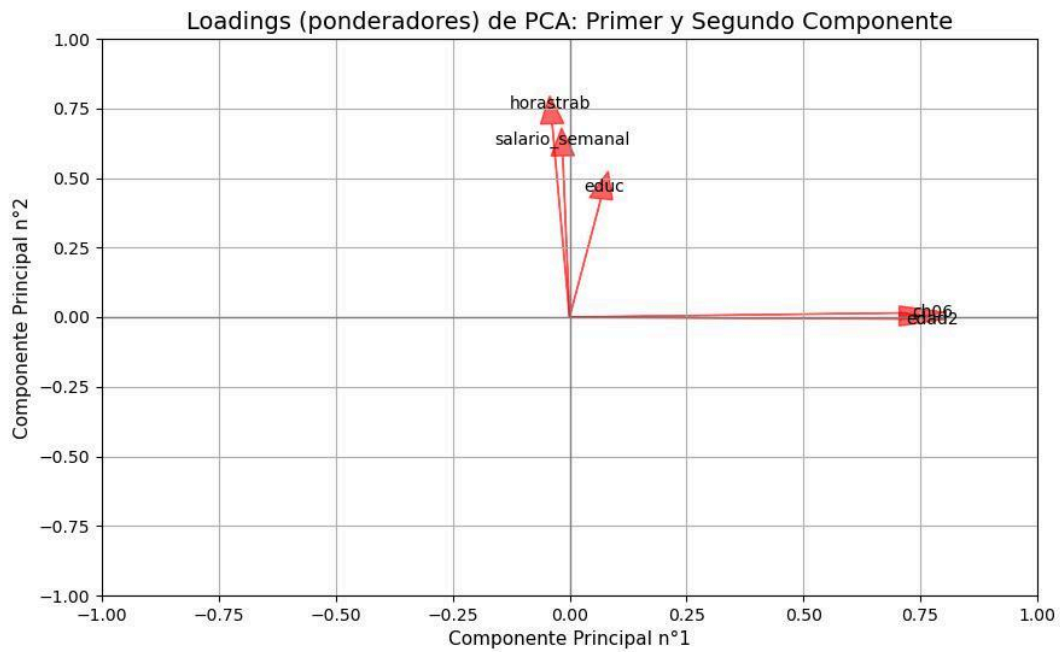
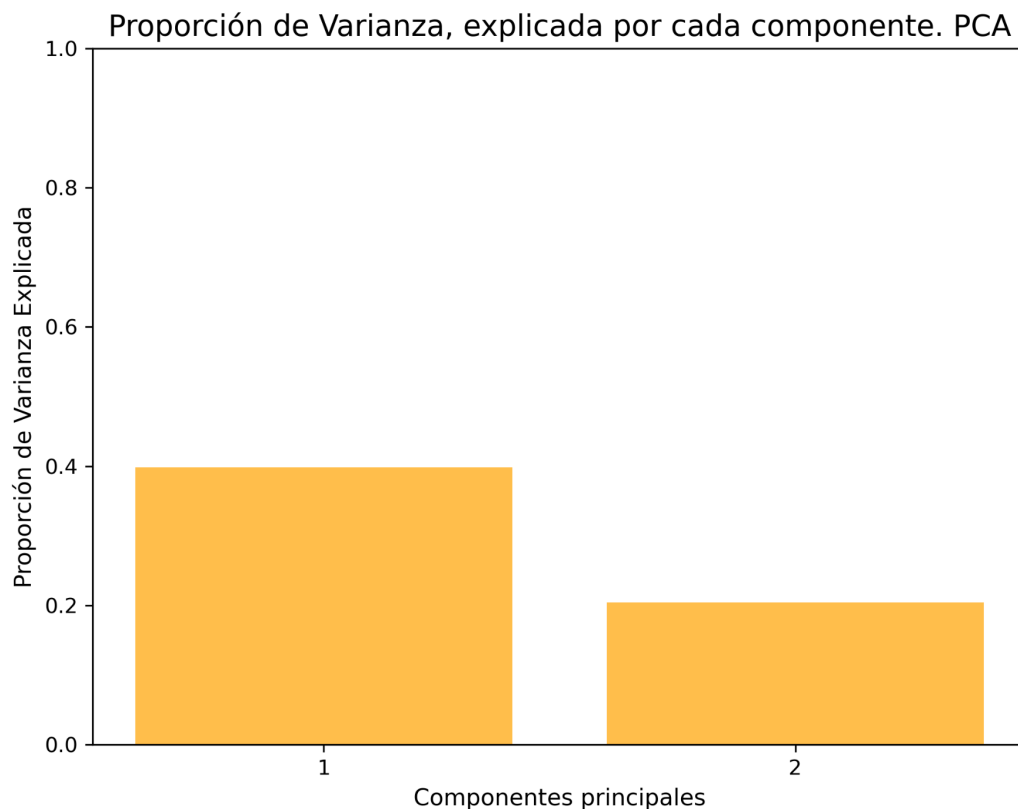


El gráfico de dispersión relaciona las observaciones existentes de los componentes principales. En el eje de las X se encuentran aquellas observaciones del primer componente, mientras que en el de las Y, las del segundo componente. Como regla general podemos decir que si los puntos se agrupan en la parte superior derecha del gráfico, tienen valores altos en ambos componentes, lo cual indicaría que las observaciones comparten atributos que las colocan en los lugares de mayor varianza de ambos componentes. Lo opuesto se podría decir si se agrupan en la parte inferior izquierda, como sucede en este caso; tenemos un conjunto denso de observaciones que se agrupa en los valores bajos de varianza de ambos componentes. Podríamos suponer que en base a la agrupación de datos en la parte inferior izquierda del dato, las observaciones comparten características similares y coherentes con una menor cantidad de horas trabajadas, un menor salario semanal o un menor nivel educativo.



En el eje X tenemos expresado el valor de los ponderadores para el primer componente principal, mientras que en el eje de las Y, el valor de los ponderadores para el segundo componente principal.

La dirección de las flechas indica su correlación positiva o negativa con el componente. Como vemos, las horas trabajadas, el salario semanal y la educación tienen una correlación positiva con el segundo componente, mientras que la edad tiene una correlación positiva con el primer componente. Por otro lado, el tamaño de la flecha indica la magnitud del loading (ponderador), siendo mayor la magnitud en el caso de la edad, pero menor en el caso del resto de variables.



El eje X muestra los dos componentes principales. El eje Y, la proporción de varianza explicada por cada componente. Los valores exactos son un 39,8% de la varianza total de los datos explicada por el primer componente, y un 20,4% explicada por el segundo componente. A pesar de que el primer componente explica mayor parte de la varianza que el segundo, lo cual indicaría que es más determinante a la hora del análisis, la explicación de la varianza que se le atribuye al primer componente no supera el 50%. Esto indica que no es un valor alto; podría afirmarse que el análisis de la varianza entonces adquiere cierta complejidad, ya que es necesario sumar dimensiones (o componentes adicionales) para explicar el total de la varianza.