1. Tanto para la base 2004 como para 2024 vemos que la diferencia entre la media para la base de entrenamiento y la base de test es relativamente baja en la mayoría de las variables. Es decir, si bien hay diferencias en los valores del promedio estas son bastante similares, por lo que podríamos decir que la base de entrenamiento es representativa de la base de testeo.
2. Tabla:

|  | Modelo 1 | Modelo 2 | Modelo 3 | Modelo 4 | Modelo 5 |
| --- | --- | --- | --- | --- | --- |
| Edad | -1117.443 | -5087.510 | -7310.923 | -6724.342 | -3362 |
| Edad2 |  | 426.321 | 653.929 | 610.846 | 610 |
| Educ |  |  | 1525.947 | 1646.343 | 1646 |
| mujer |  |  |  | -2218.556 | -3703 |
| Analfabetismo |  |  |  |  | -3362 |
| Nivel Ed |  |  |  |  | 823 |
| observaciones | 3224 | 3224 | 3224 | 3224 | 3224 |
| R^2 | 0 | 0 | 0.04 | 0.052 | 0.0516 |

Podemos ver que tanto los modelos 1 y 2 tienen un R^2 de 0. A partir de la variable Educación (entendido como máximo nivel educativo alcanzado, CH12), la cual tiene una relación positiva con el salario semanal aumenta R^2. Un mayor se ve al añadirse la variable de género (en este casó mujer), teniendo esta una relación negativa con respecto al salario semanal. Lo mismo ocurre con una de las variables escogidas (analfabetismo), esto significa que al aumentar este, el salario semanal disminuye.

Y por último el modelo 5, el más completo incluye la variable Nivel educativo (entendido esta vez como máximo nivel educativo completado), esto se distingue de educación (como se veía antes) ya que esta toma en cuenta simplemente los niveles educativos terminados, aunque esta diferencia parezca menor, podemos ver como al agregar esta variable y la consideración del analfabetismo, el modelo 5 toma mucha más significancia.

Cabe aclarar que en todos los modelos el número de observaciones fue el mismo y se puede ver una evolución progresiva, modelo a modelo, en el R^2.