

Proyecto Modelos Lineales

Oscar Gamboa, Gonzalo Mardones, Nicolas Montecinos

2022-10-13

Introducción

- bla bla

Planteamiento del Problema

- Fenomeno: Licencias medicas en funcionarios de educación en Huasco - 2022 han llegado con mucha frecuencia...

Justificación del Problema

Las razones que conducen a investigar el fenómeno.

- En funcion del impacto de las licencias medicas. - caracterizacion de la licencia medica (numeros promedios de días)

Estado del Arte

BUSCAR UNA REFERENCIA A NIVEL NACIONAL PARA DECLARAR EL IMPACTO (DOCENTES Y MEDICOS) METODOLOGIA DE CONTEO (DESCRIPTIVO)

Objetivo General y Objetivo Especifico

- Obj General: se busca caracterizar la catnidad de dias promedios de LM de los docentes contratados bajo la administración del servicio de ed. publica de la provincia de Huasco, con info actualizada ENE-2022 a OCT-2022 con el fin poder entregar mayores antecedentes a la hora de tomar decisiones a nivel provincial
- Obj Especificos:
 - Obtener, recopilar y extraer las características más importantes que aporten al estudio de la investigación
 - Descripción y asociación de variables predictoras con respecto a la cantidad promedio de días.
 - validación la base de datos y modificar su estructura para los fines de este estudio

Plateamiento de la hipótesis en el estudio

- Se quiere determinar si la evaluación docente esta indirectametne relacionada con el promedio de días tomados por licencias medicas de los mismos docentes

Generación de unidades de información

Diseño de estrategias metodológicas

Debido a lo que se quiere explicar, se plantea metodologia a traves de regresiones lineales debido a que se quiere explocar

Procesamiento de las unidades de información

Descripción de los datos

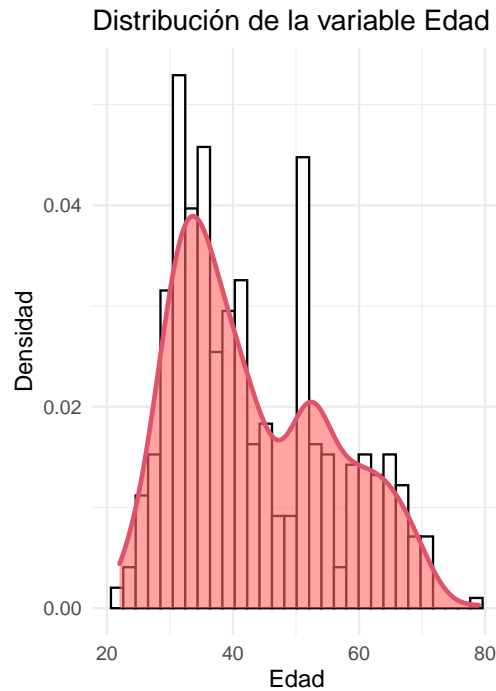
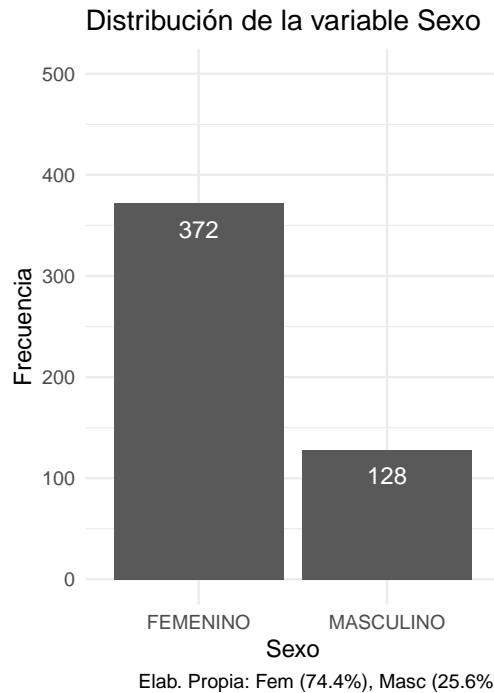
Base de datos de Licencias Médicas		
Variable	Tipo	Descripción
SEXO	Factor	Masculino y Femenino
ESTADO_CIVIL	Factor	Casado(a), Viudo(a), Soltero(a), Divorciado(a)
EDAD	Double	Edad
RENTA_PROMEDIO	Double	Renta Promedio en Pesos Chilenos
SISTEMA_SALUD	Factor	Fonasa, Isapre
JORNADA	Double	Mediadas en horas semanales
NIVEL	Factor	Básica o Media
TIPO_ESTABLECIMIENTO	Factor	Escuela o Liceo
CALIDAD_DESEMPEÑO	Factor	Profesor Titular o no
TRAMO_DOCENTE	Factor	Inicial, Temprano, Acceso, Avanzado, Experto I, Experto II, Sin Encasillamiento
EVALUACION_DOCENTE	Factor	Sin Evaluación, Destacado, Básico, Insatisfactorio, Competente
PROMEDIO_DIAS_LM	Double	Promedio Días de Licencias Médicas
TRASLADO_COMUNA	Double	SI (1) o NO (0)

Estadística Descriptiva

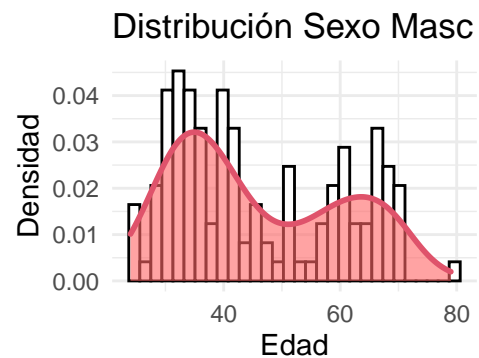
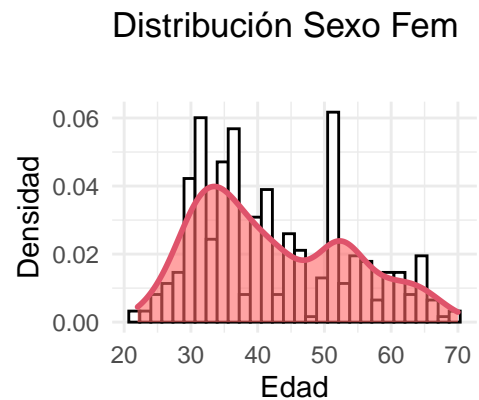
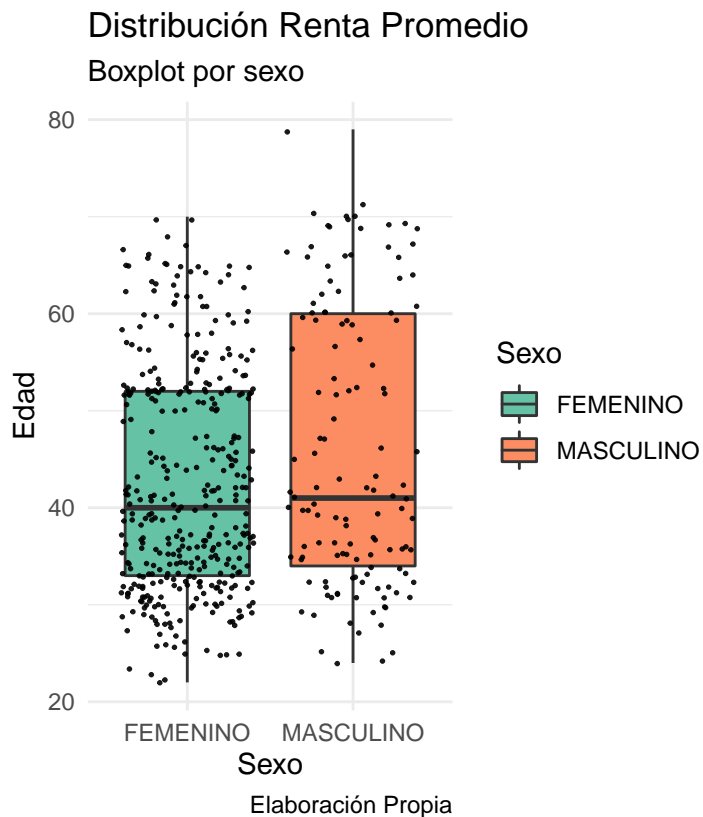
A continuación se presenta análisis descriptivos para las variables más relevantes del estudio, de las cuales destacamos días promedios de licencias médicas,

Distribución de Licencias Médicas

El gráfico de la izquierda muestra la distribución de licencias médicas que incluyen a los docentes/funcionarios que no han presentado licencias médicas hasta la fecha, en el gráfico de la derecha se muestra la distribución de licencias médicas pero descartando los casos de docentes/funcionarios que no han presentado licencias médicas, siendo este el grupo de personas empleadas para el análisis

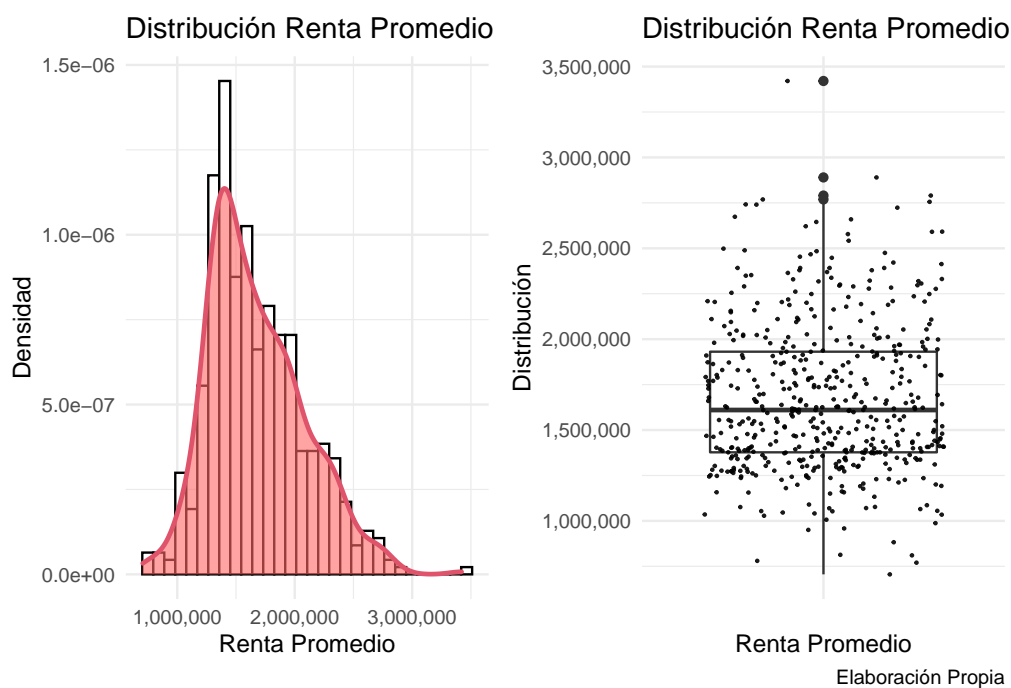


De acuerdo a los resultados, la variable edad parece tener una distribución bimodal, de hecho, nos da luces de una cierta mezcla de distribuciones, por lo cual se hace necesario el poder analizar esta variable de acuerdo a variables categóricas como es el sexo.



En este sentido, la distribución de edades en mujeres y hombres es bastante similar, en cada sexo se puede apreciar una distribución bimodal.

Min.	1st Qu.	Mediana	Promedio	3rd Qu.	Max.
\$705.804	\$1.377.071	\$1.609.868	\$1.668.484	\$1.931.119	\$3.420.643



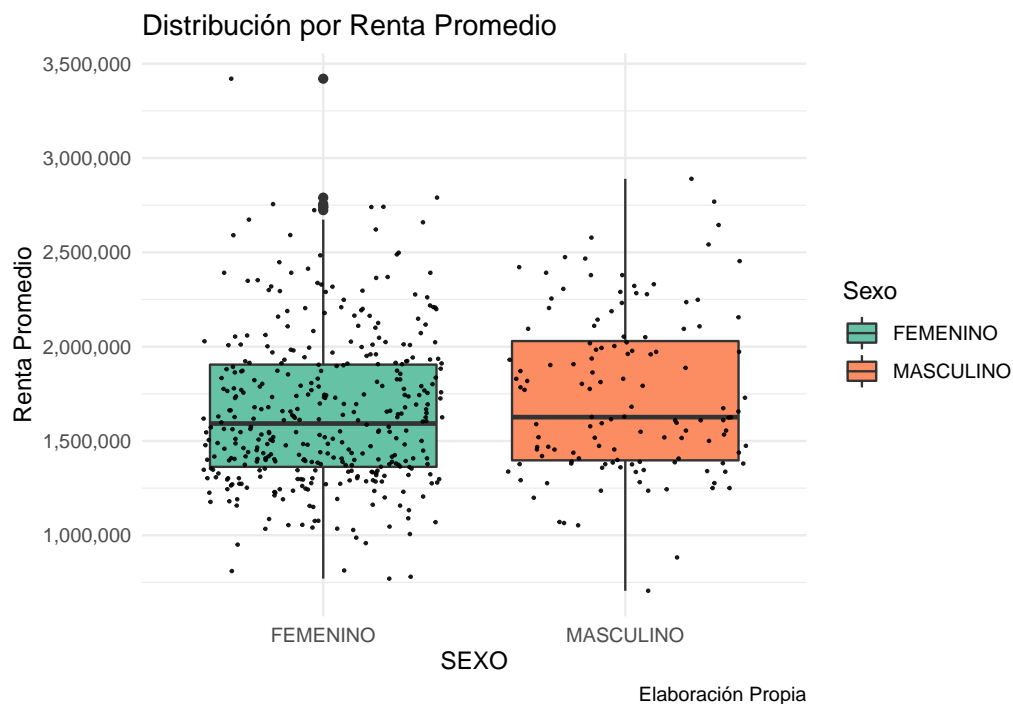
En este caso la distribución de la renta está centrada alrededor de 1,7 millones de pesos, donde el 25% más bajo de la distribución de renta está por debajo de los 1.37 millones, mientras que el 25% de los casos con mayores rentas se encuentran entre 1.93 y 3.42 millones de pesos.

Femenino

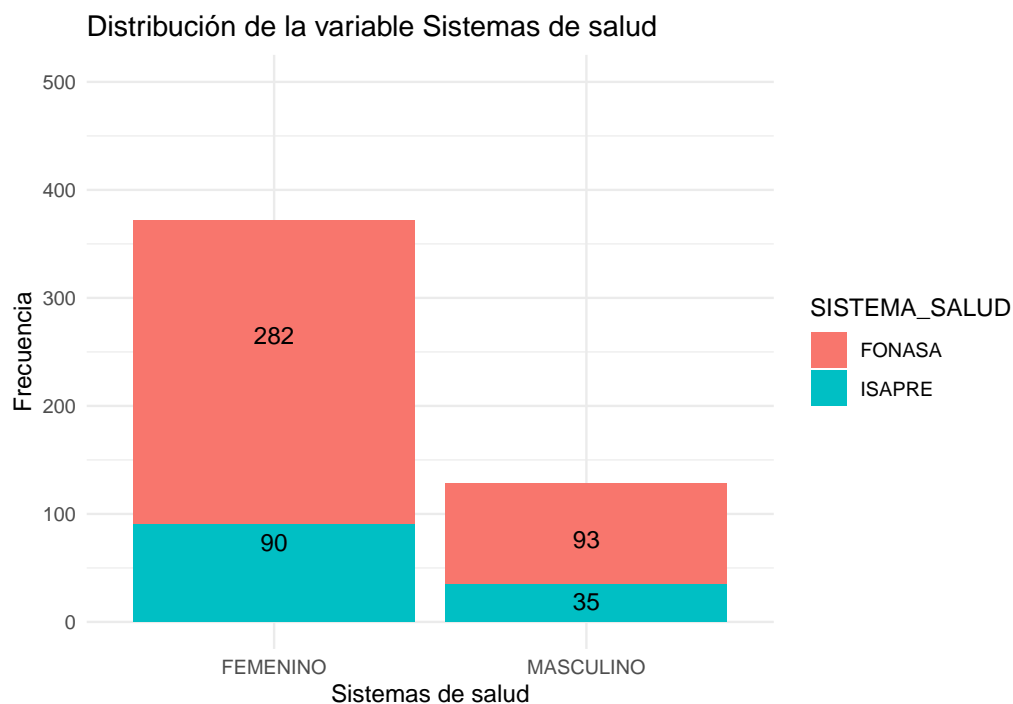
Min.	1st Qu.	Mediana	Promedio	3rd Qu.	Max.
\$770.197	\$1.362.827	\$1.592.526	\$1.652.053	\$1.904.939	\$3.420.643

Masculino

Min.	1st Qu.	Mediana	Promedio	3rd Qu.	Max.
\$705.804	\$1.397.245	\$1.626.350	\$1.741.390	\$2.029.702	\$3.233.292



Comparando los niveles de renta por sexo se puede notar una cierta similitud en el rango de montos promedio, más aún, las estadísticas de posición (cuartiles) son bastante similares, por lo que podría no haber un efecto del sexo sobre los ingresos promedio. Para las mujeres la mediana de las rentas promedio es 1.58 millones, mientras que para los hombre bordea los 1.62 millones de pesos.

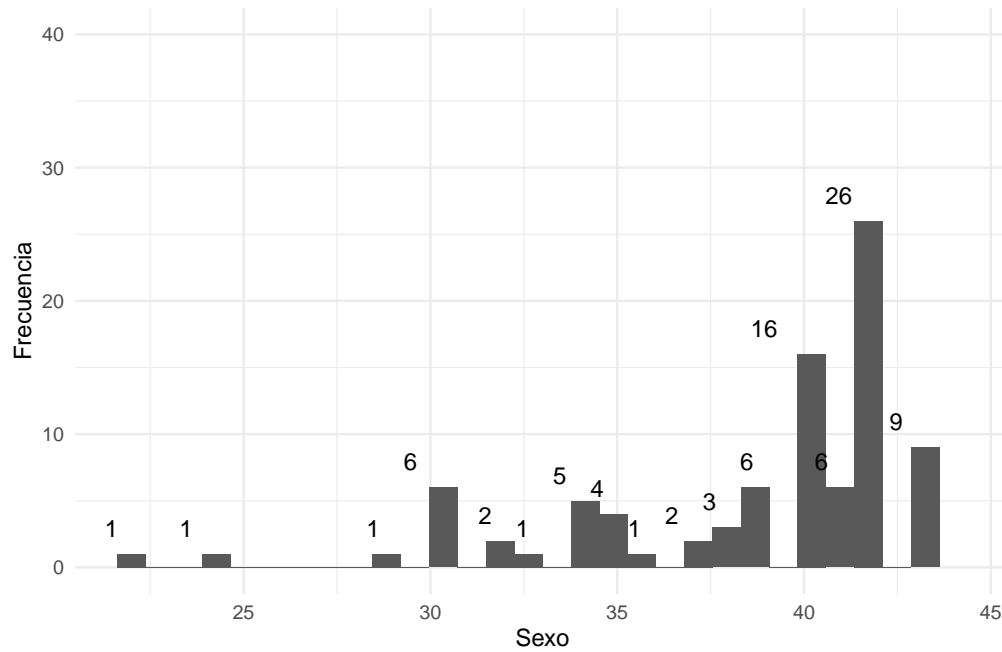


Según los datos, la proporción de personas que pertenecen al sistema de salud fonasa es alto tanto en hombres como en mujeres, esto tiene sentido desde el punto de vista que es una base de información de la provincia de Huasco donde existe una alta mayoría de zonas rurales donde el sistema fonasa es predominante.

Min. 1st Qu. Median Mean 3rd Qu. Max.

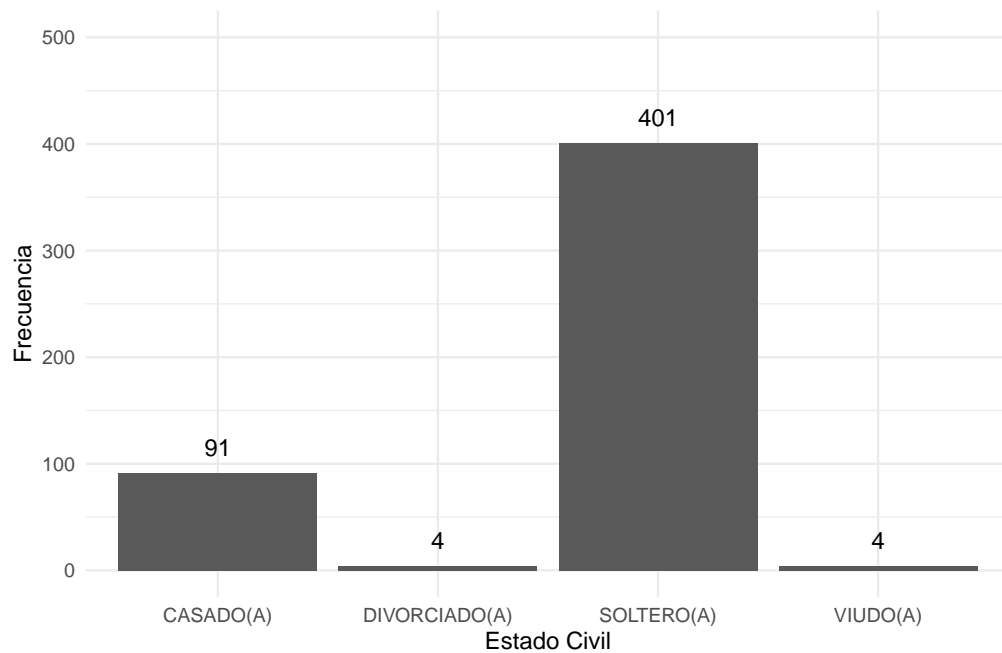
22.00 44.00 44.00 43.05 44.00 44.00

Distribución de Jornada por Sexo



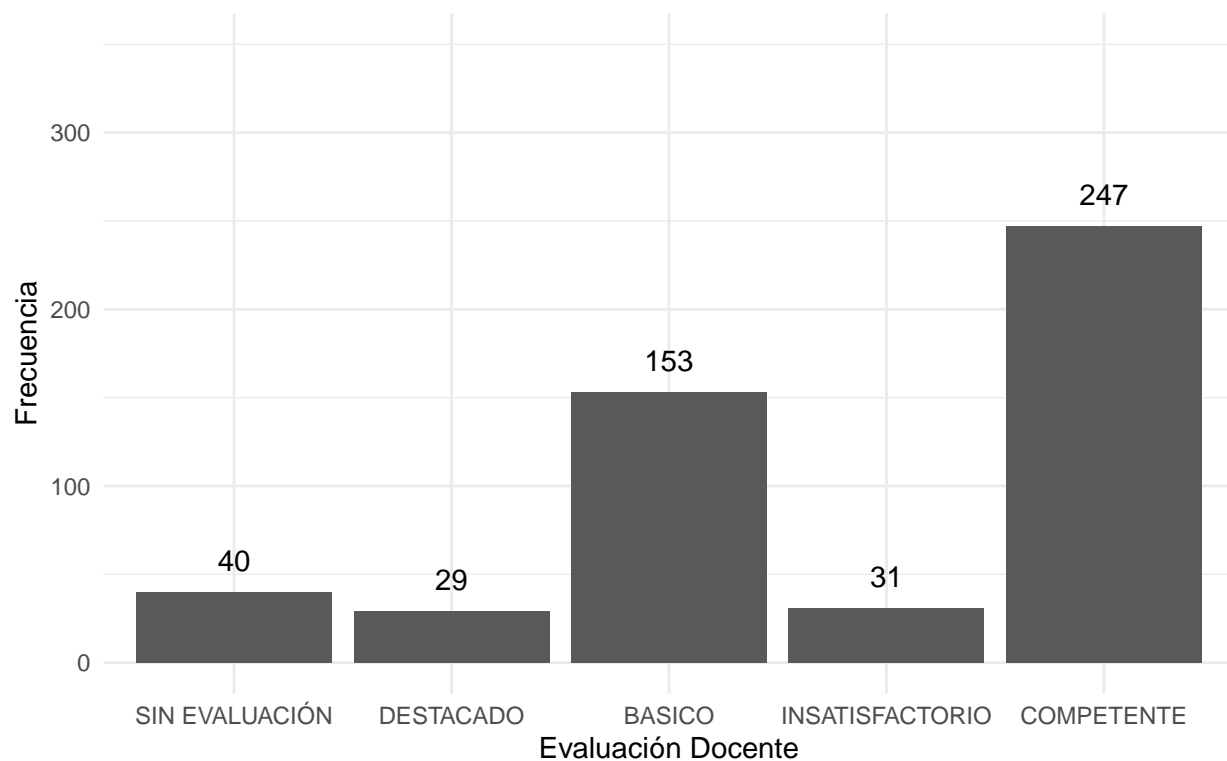
Podemos notar que prácticamente el 75% de las personas de la muestra trabajan jornadas de 44 o menos horas semanales, mientras que sólo un caso muestra 88 horas laborales semanales, este caso se debe a (justificar el caso)

Distribución de Estado Civil



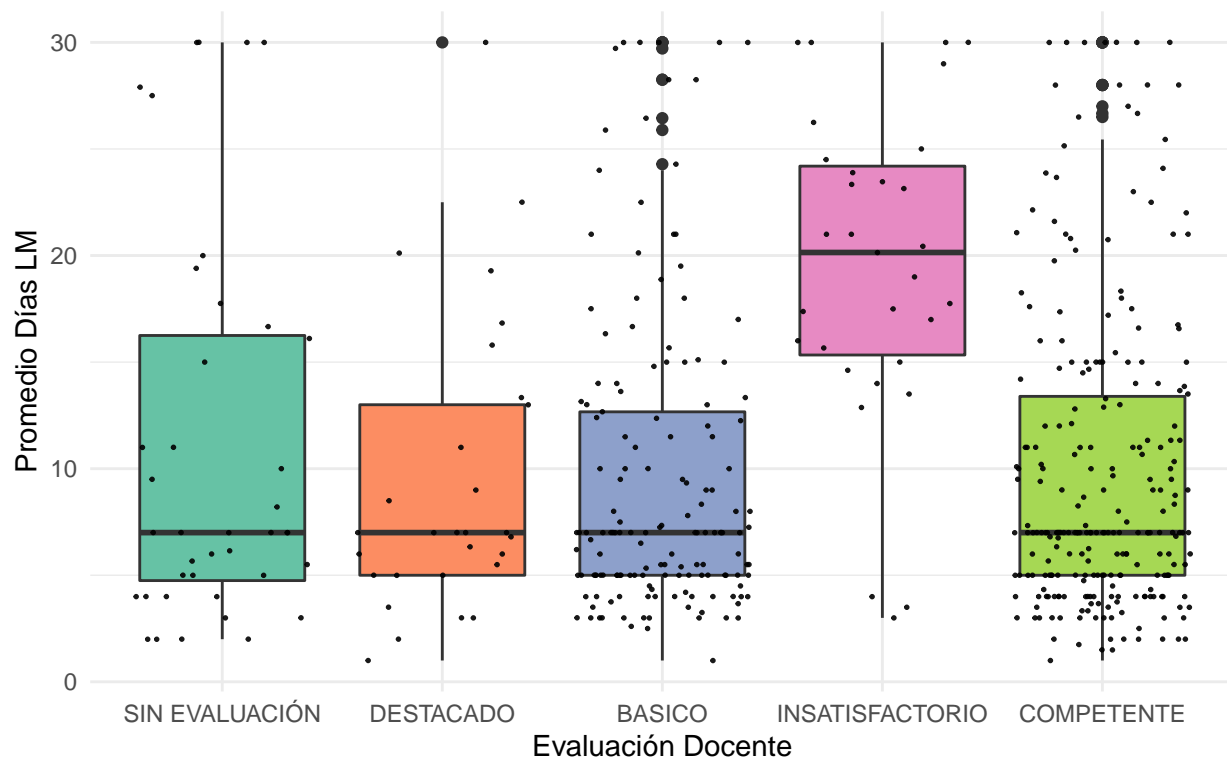
texto

Distribución de Evaluación Docente



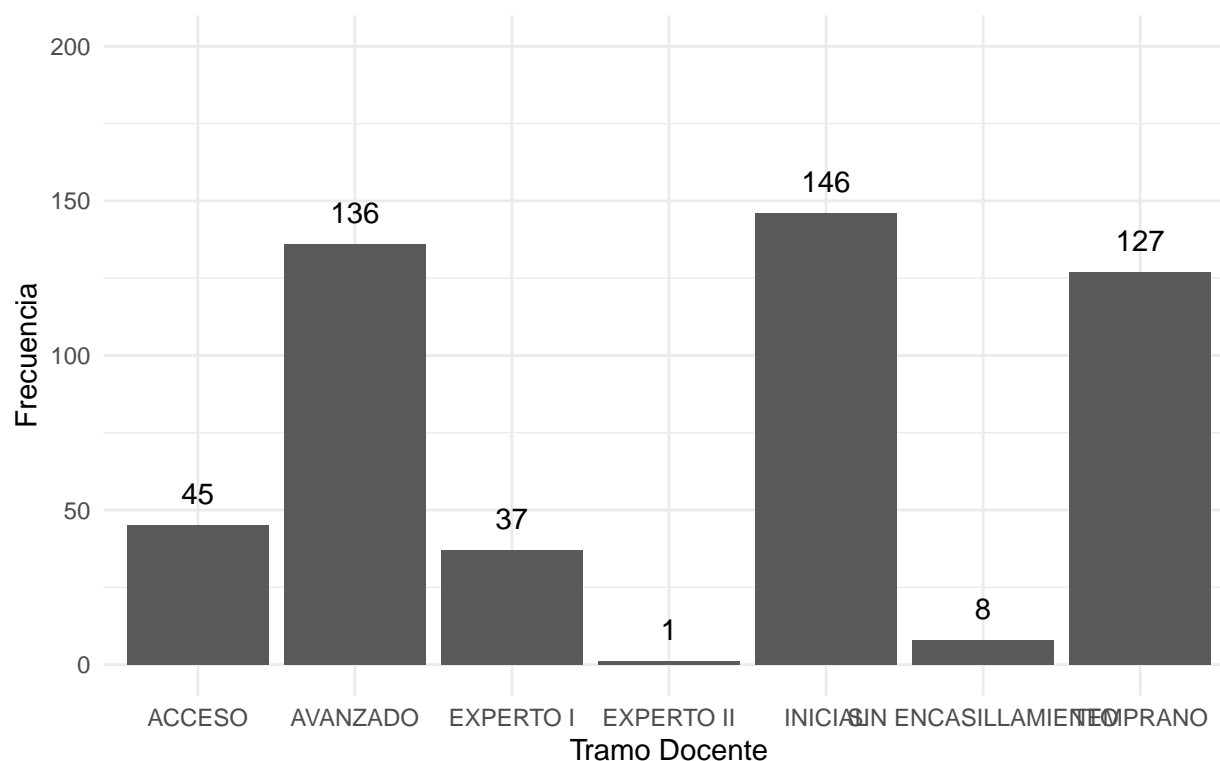
Elaboración Propia

Distribución de Evaluación Docente



Elaboración Propia

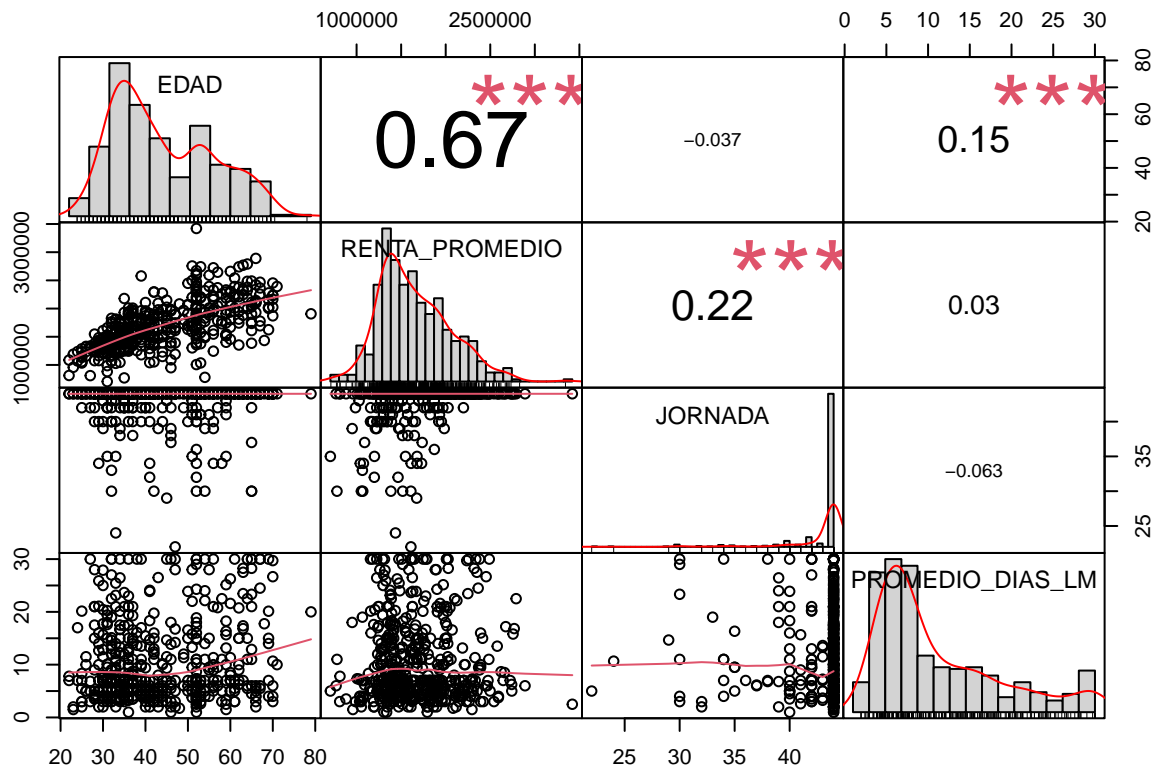
Distribución de la variable Tramo Docente



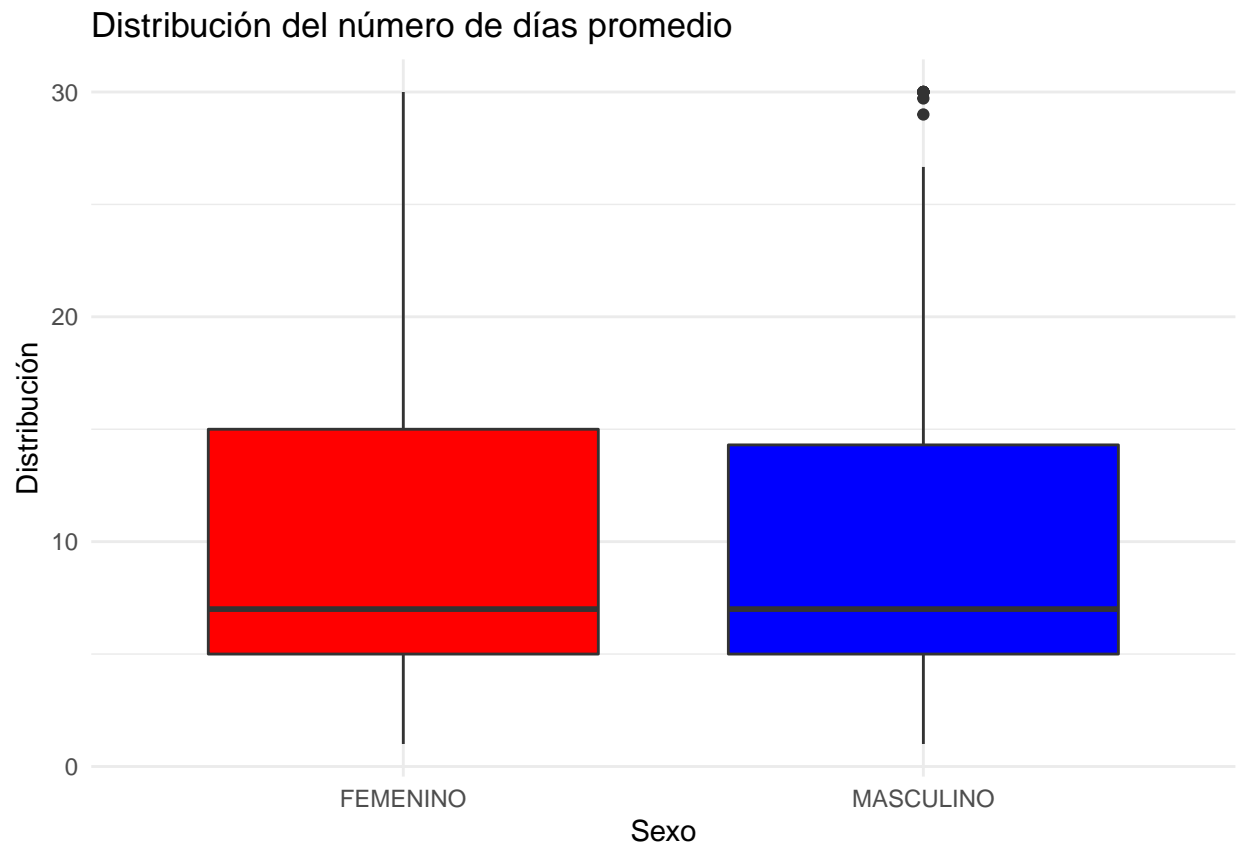
Elaboración Propia

Correlación

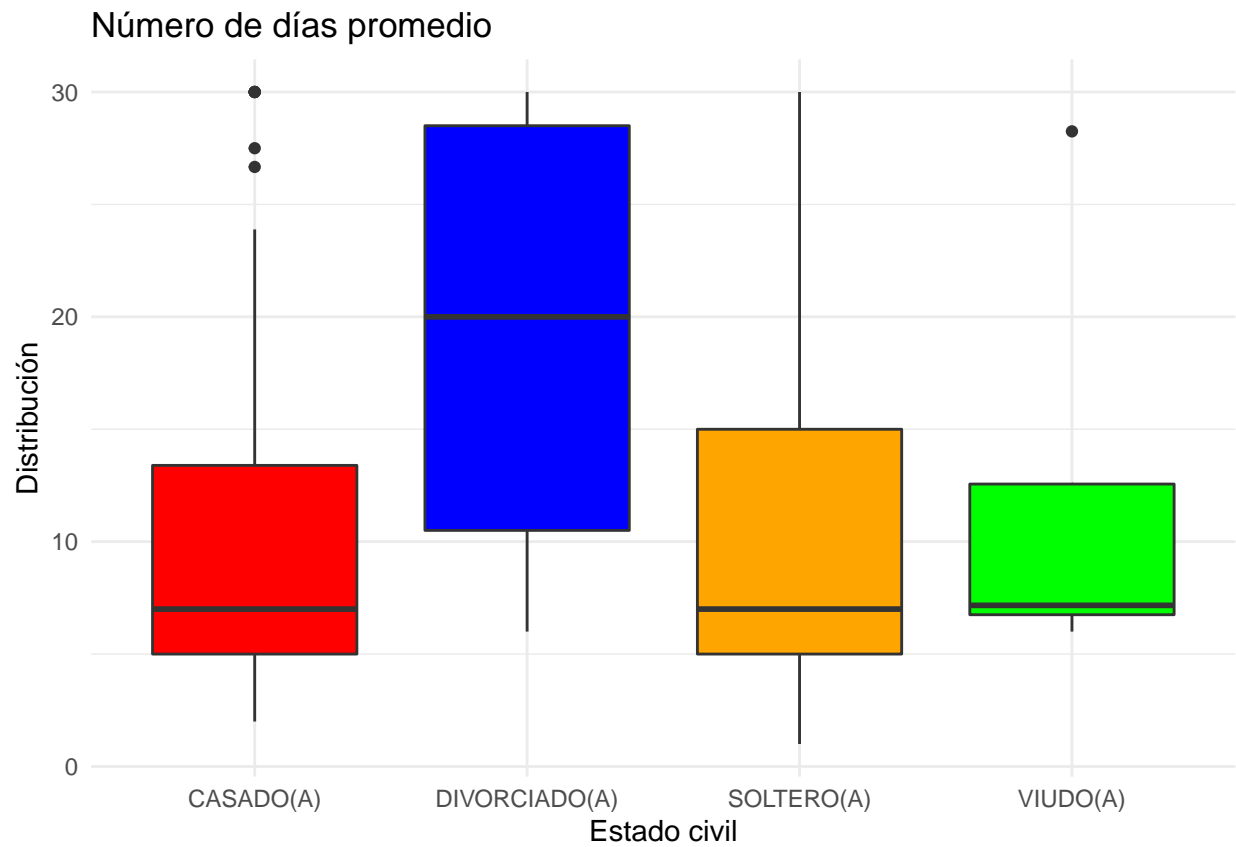
```
## # A tibble: 6 x 24
##   ESCALAFON RUT_FUNCIONARIO NOMBRE      AP_P AP_M SEXO ESTADO_CIVIL EDAD
##   <chr>      <chr>          <chr>    <chr> <chr> <fct> <fct>      <dbl>
## 1 DOCENTE  16184521-3    MARCELA ALEJAN~ MEND~ GALL~ FEME~ SOLTERO(A)    36
## 2 DOCENTE  12940875-8    ROBERTO CESAR  GOYA  TORR~ MASC~ CASADO(A)    46
## 3 DOCENTE  8240042-7     NORMA ISABEL  COVA~ DIAZ  FEME~ SOLTERO(A)    64
## 4 DOCENTE  16111923-7    CLAUDIA STEPHA~ OJEDA ANGU~ FEME~ SOLTERO(A)    37
## 5 DOCENTE  9267343-K     TATIANA DE LOU~ LEDE~ ROJAS FEME~ CASADO(A)    61
## 6 DOCENTE  7680751-5     ESTELA EDIT   VELA~ ALVA~ FEME~ SOLTERO(A)    63
## # ... with 16 more variables: COMUNA_RESIDENCIA <chr>, RENTA_PROMEDIO <dbl>,
## #   SISTEMA_SALUD <fct>, JORNADA <dbl>, NIVEL <fct>, ESTABLECIMIENTO <chr>,
## #   TIPO_ESTABLECIMIENTO <fct>, COMUNA_ESTABLECIMIENTO <fct>,
## #   CALIDAD_DESEMPEÑO <fct>, ESTAMENTO <fct>, TRAMO_DOCENTE <fct>,
## #   EVALUACION_DOCENTE <fct>, CANTIDAD_LM <dbl>, PROMEDIO_DIAS_LM <dbl>,
## #   SUMA_DIAS <dbl>, TRASLADO_COMUNA <fct>
```

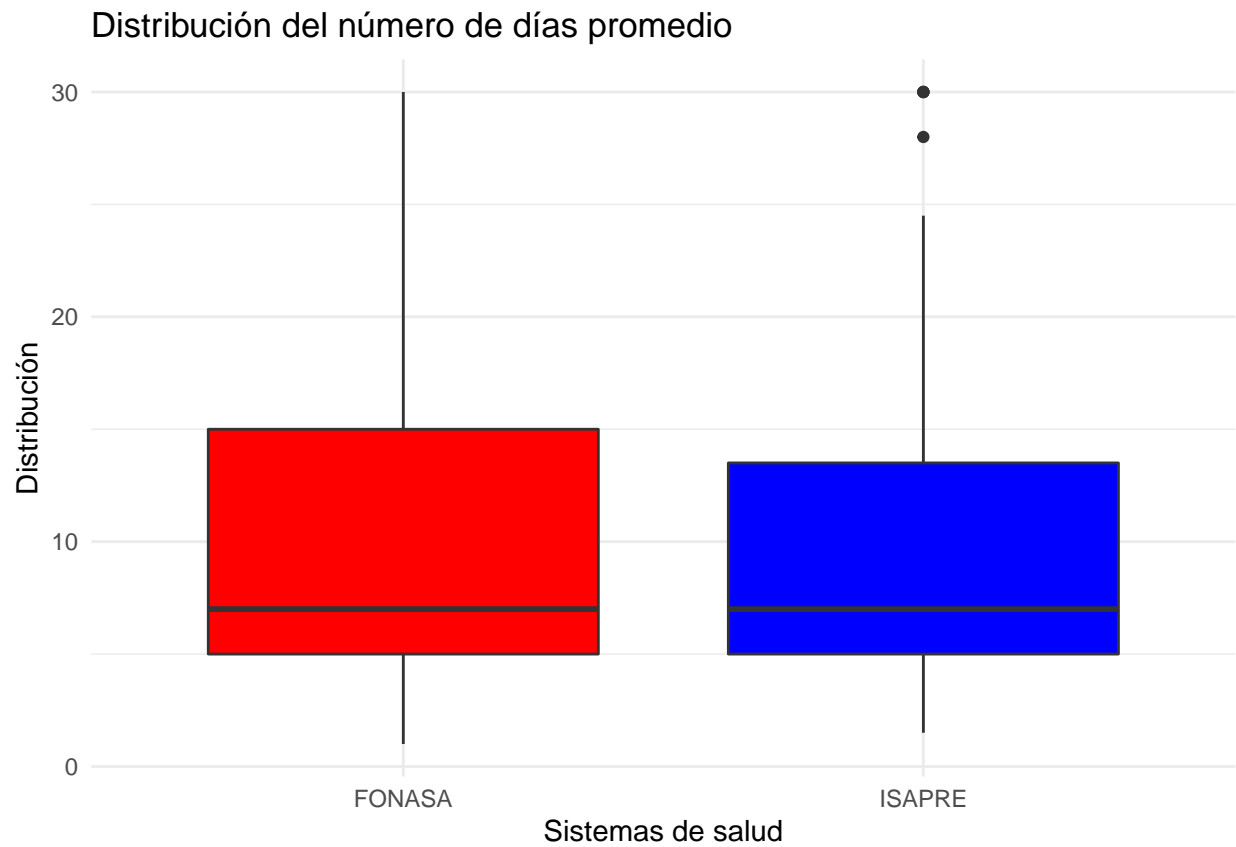
Al analizar las correlaciones entre las variables predictoras continuas y el número promedio de días con licencia se puede apreciar que la mayor correlación la presenta la variable edad con un coeficiente de 0.15 lo que sugiere posible efecto explicativo, seguido por la renta promedio con una correlación de 0.043, mientras que la jornada presente una correlación menor y negativa de -0.063.



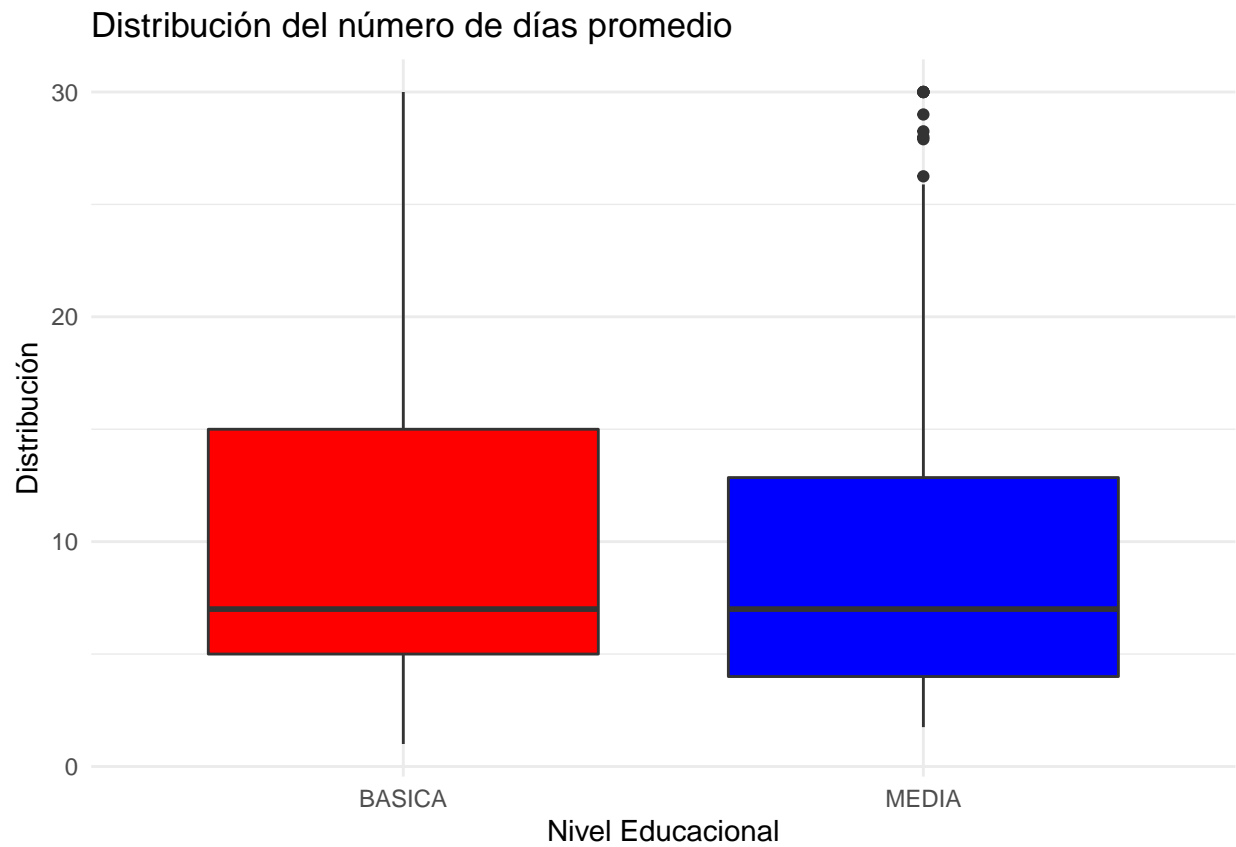
En este caso, los boxplot no sugieren un efecto del sexo sobre la distribución del numero de días con licencia promedio, dado que ambos soportes y estadísticas de posición son similares.



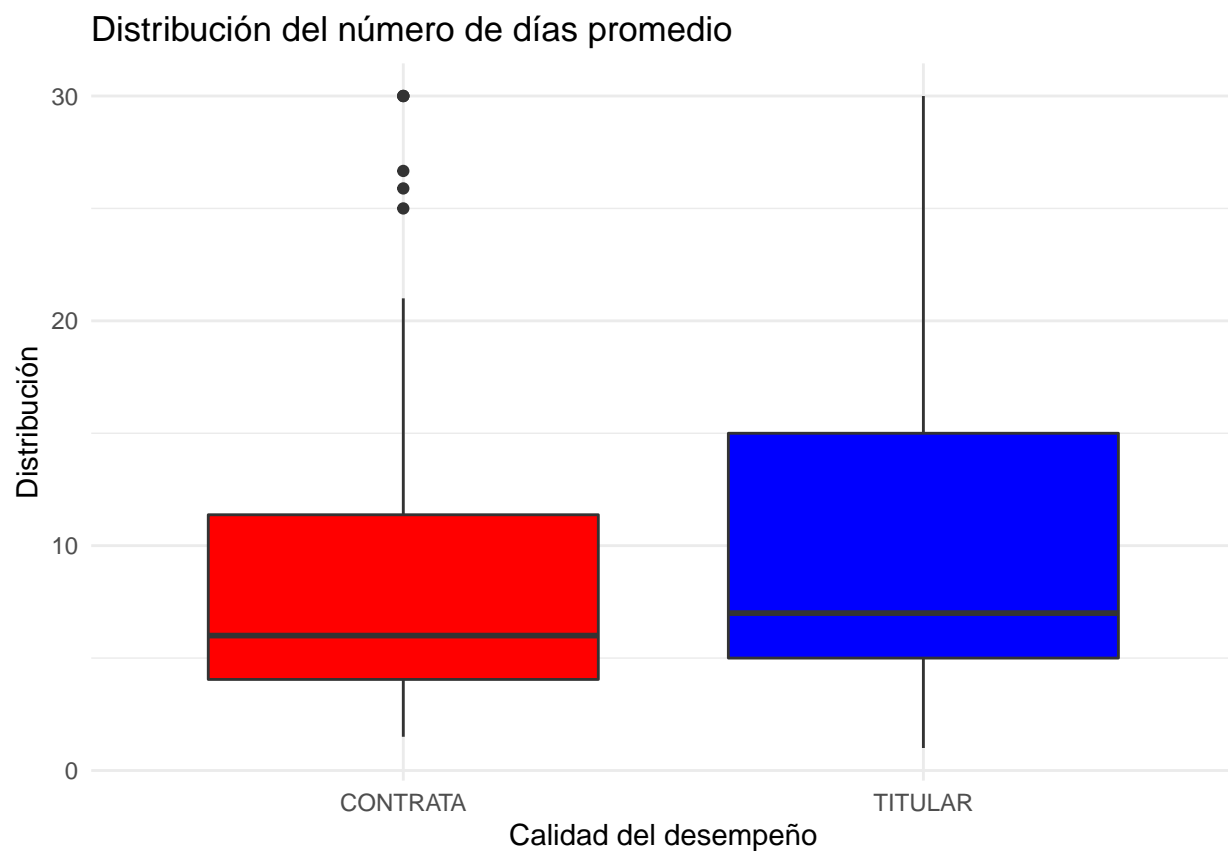
De acuerdo al gráfico los datos sugieren que el nivel “Divorciado” podría tener un efecto en el número promedio de días con licencia, dado que su soporte es levemente superior al de los demás niveles.



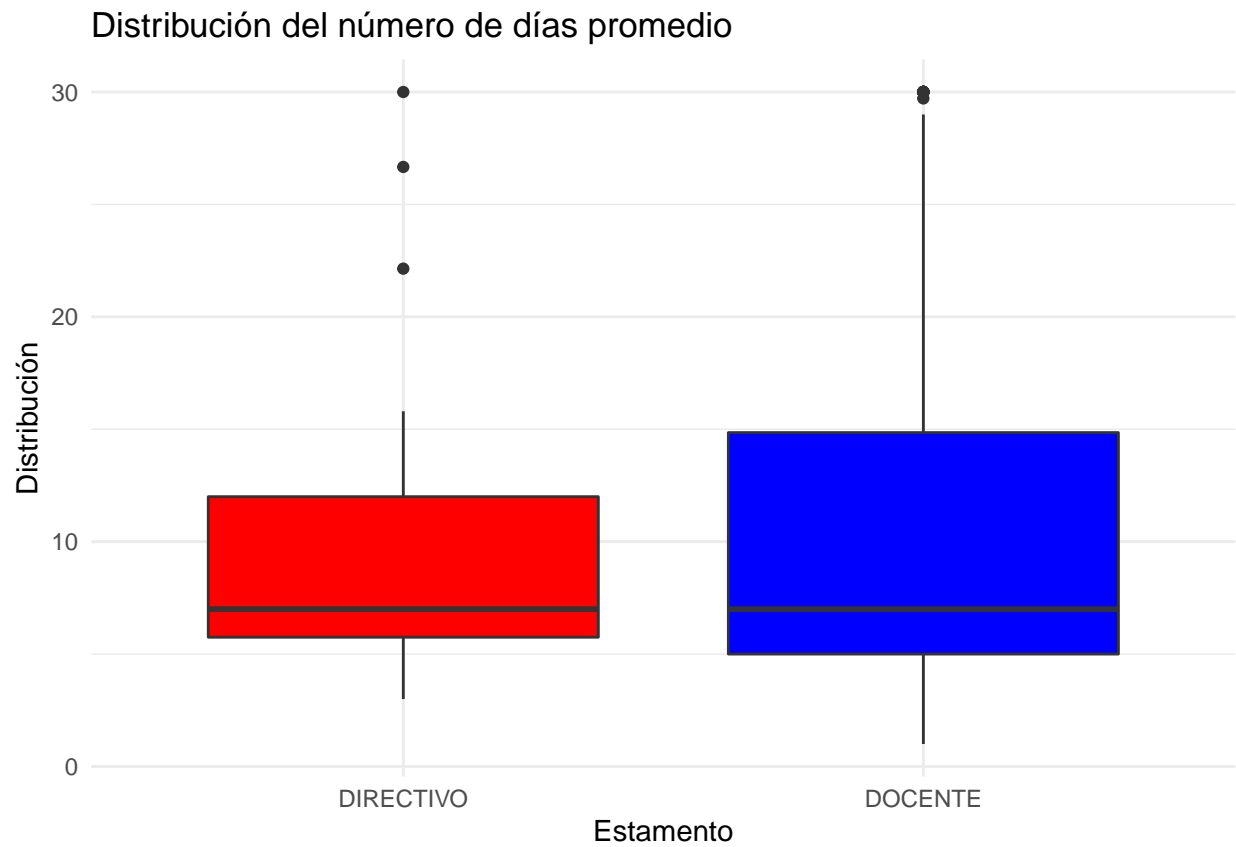
Para el caso de los sistemas de salud, no se observan diferencias importantes en los soportes ni en las estadísticas de posición del número promedio de días con licencia, es decir, el sistema de salud podría no tener un efecto significativo.



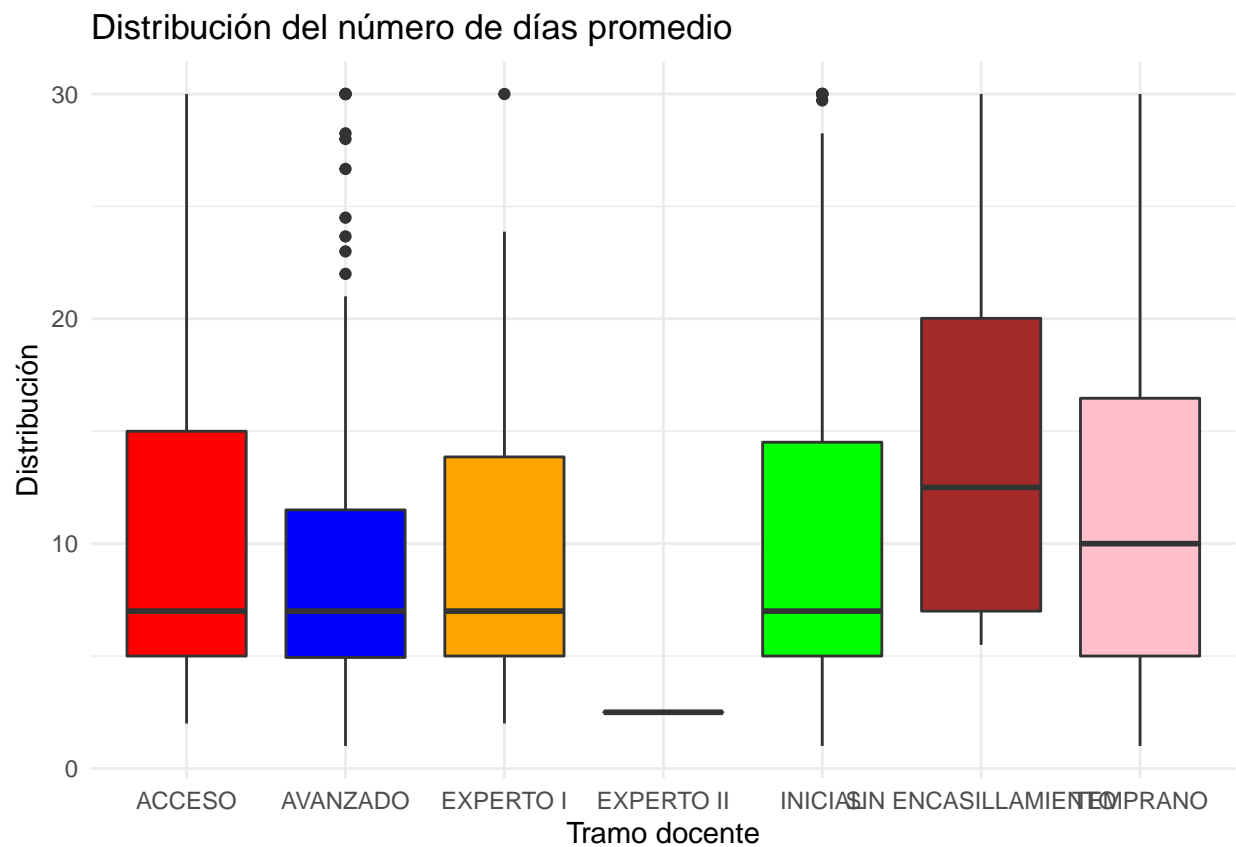
Al observar la distribución del número promedio de días con licencia por nivel educativo no se observan grandes diferencias, lo que sugiere que este factor no presenta un grado de asociación significativo.



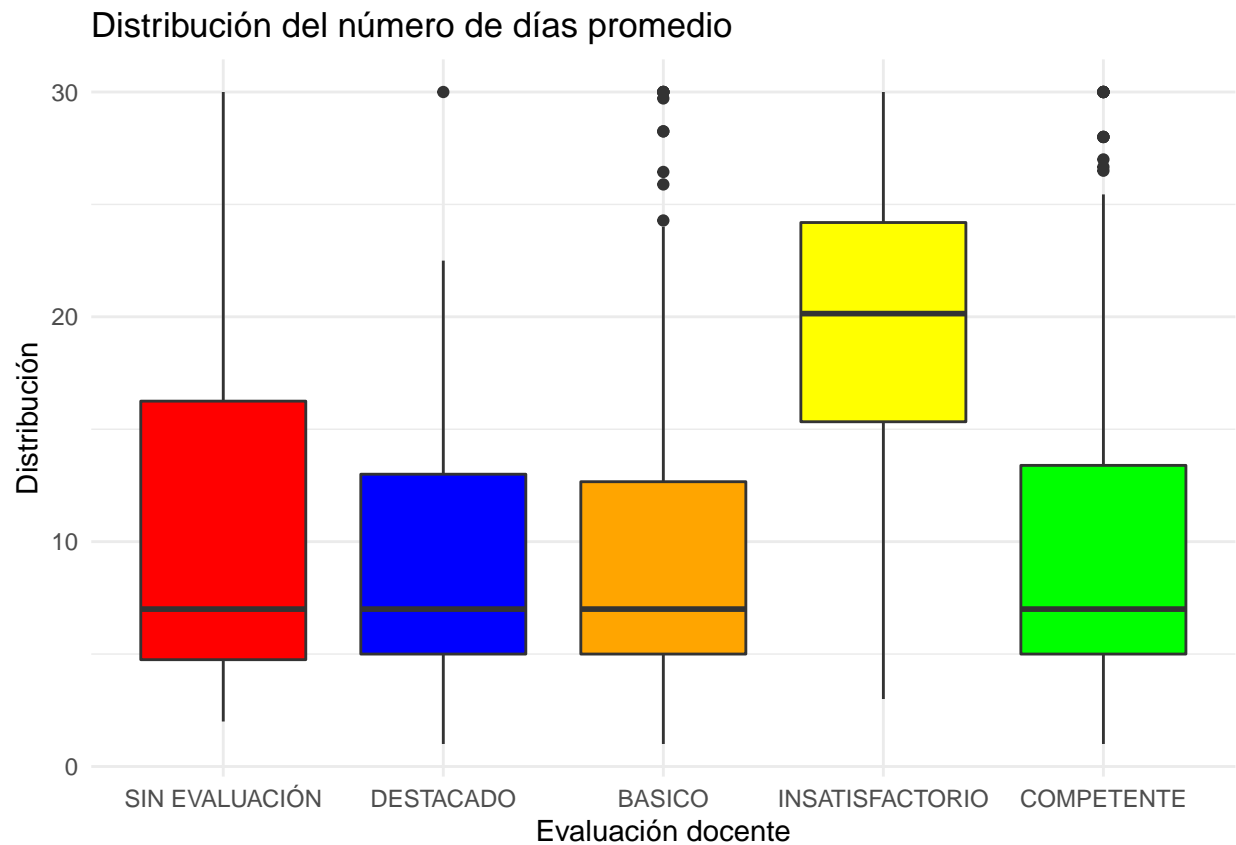
En este caso, el ser docente titular presenta un soporte levemente mayor que aquel docente con modalidad contrata, lo que puede sugerir un grado de asociación de este factor con el número promedio de días con licencia.



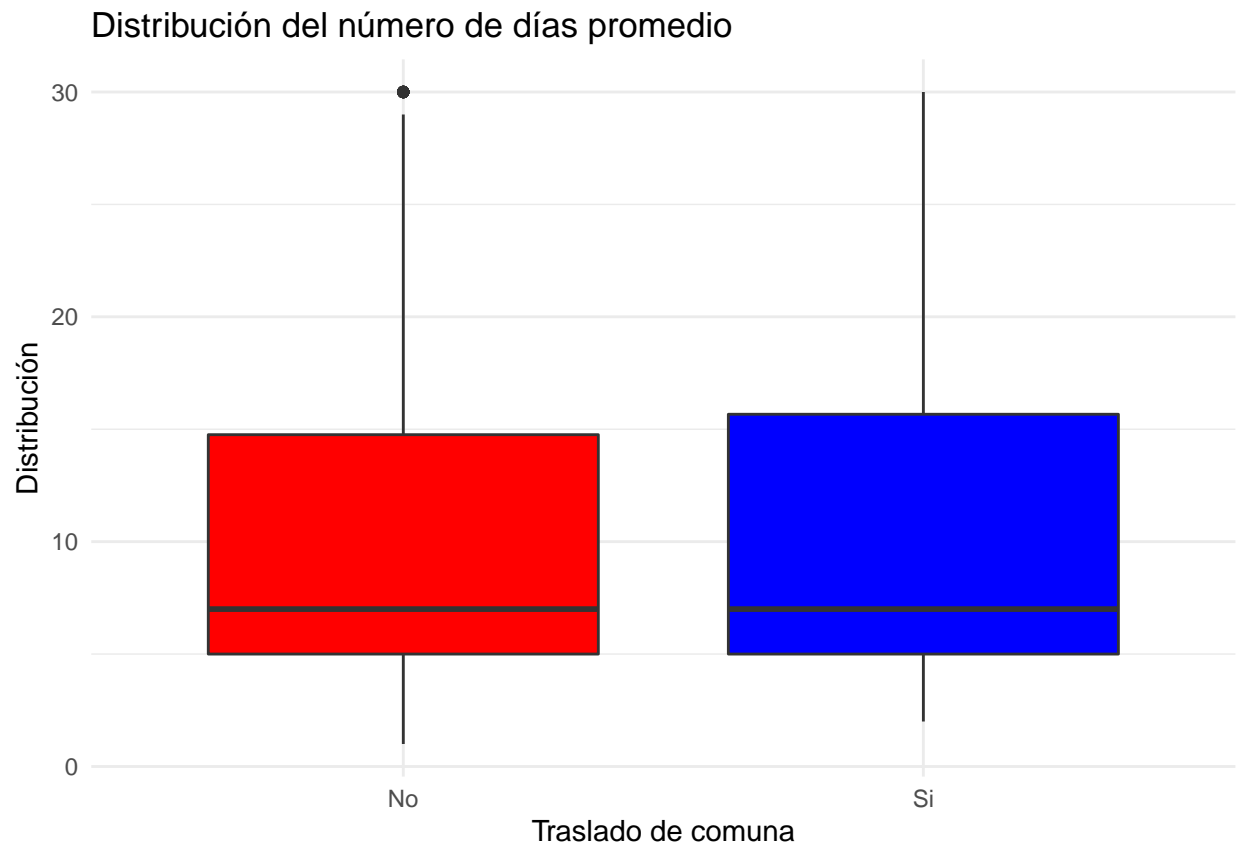
Analizando el factor “Estamento” se puede notar que aquel docente que sólo presneta estamento de docente tiene mayor soporte que aquel que tiene estamento de directivo, lo que sugiere algún grado de asociación de este factor con el número promedio de días de licencias.



En este caso los boxplot de los días promedio de licencias analizados por los niveles del tramo docente muestran cambios en sus soportes lo que puede sugerir algún asociación de este factor.



De acuerdo al gráfico se aprecia una distribución de valores centrales diferente en el nivel “insatisfactorio” del factor “Evaluación docente”, mientras que el resto de los niveles presentan soportes y distribuciones centrales similares, esto sugiere que se podría crear una variable dummy que tome valor uno cuando la evaluación docente es insatisfactorio o cero en otro caso, pero se tendría que evaluar si produce los mismos asociaciones sobre el modelo que se proponga.



Respecto a la distribución del número de días con licencia de acuerdo al traslado, los soportes y distribuciones son similares, por lo cual no se sugiere alguna asociación de este factor sobre la variable respuesta (número de días promedio con licencias)

Selección Formal de Modelo

Forward

Para la selección del modelo forward se utilizó una significancia del 5% en cada uno de los tests de hipótesis realizados

```
modelo0 = lm(PROMEDIO_DIAS_LM ~ 1, data = BBDD_proyecto_sin_ceros)
add1(modelo0, ~ . + SEXO + ESTADO_CIVIL + EDAD + RENTA_PROMEDIO +
      SISTEMA_SALUD + JORNADA + NIVEL + CALIDAD_DESEMPEÑO +
      ESTAMENTO + EVALUACION_DOCENTE + TRAMO_DOCENTE + TRASLADO_COMUNA,
      test="F")
```

```
## Single term additions
##
## Model:
## PROMEDIO_DIAS_LM ~ 1
##
```

	Df	Sum of Sq	RSS	AIC	F value	Pr(>F)
<none>			29400	2039.1		
SEXO	1	4.29	29396	2041.0	0.0727	0.7875728
ESTADO_CIVIL	3	324.01	29076	2039.5	1.8424	0.1385146
EDAD	1	701.41	28699	2029.0	12.1712	0.0005284 ***
RENTA_PROMEDIO	1	26.17	29374	2040.6	0.4437	0.5056665
SISTEMA_SALUD	1	135.72	29265	2038.8	2.3095	0.1292163
JORNADA	1	116.36	29284	2039.1	1.9789	0.1601349

```
## NIVEL      1      232.20 29168 2037.1  3.9645 0.0470156 *
## CALIDAD_DESEMPEÑO 1      423.14 28977 2033.8  7.2720 0.0072411 **
## ESTAMENTO  1         0.42 29400 2041.1  0.0072 0.9325123
## EVALUACION_DOCENTE 4      2732.44 26668 1998.3 12.6796 7.858e-10 ***
## TRAMO_DOCENTE 6        624.90 28776 2040.3  1.7844 0.1003511
## TRASLADO_COMUNA  1         8.02 29392 2040.9  0.1358 0.7126186
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Para decidir la entrada del segundo predictor, se debe evaluar todos los modelos que ya contienen al promedio de días de licencias medicas, a lo que se agrega cada uno de los predictores restantes por separado.

De acuerdo al caso anterior, el menor valor-p corresponde al modelo `PROMEDIO_DIAS_LM ~ EVALUACION_DOCENTE`

Coeficientes	Estimación	Std. Error	t value	Pr(>t)
Intercepto	7.59163	1.49024	5.094	4.6e-07
EVALUACION_DOCENTEDESTACADO	-2.69256	1.55890	-1.727	0.08460
EVALUACION_DOCENTEBASICO	-2.08287	1.13656	-1.833	0.06732
EVALUACION_DOCENTEINSATISFACTORIO	7.87294	1.52249	5.171	3.1e-07
EVALUACION_DOCENTECOMPETENTE	-1.91725	1.12365	-1.706	0.08844
CALIDAD_DESEMPEÑOTITULAR	2.65846	0.72190	3.683	0.00025
NIVELMEDIA	-1.84583	0.63467	-2.908	0.00376
EDAD	0.06085	0.02425	2.509	0.01233

Backward