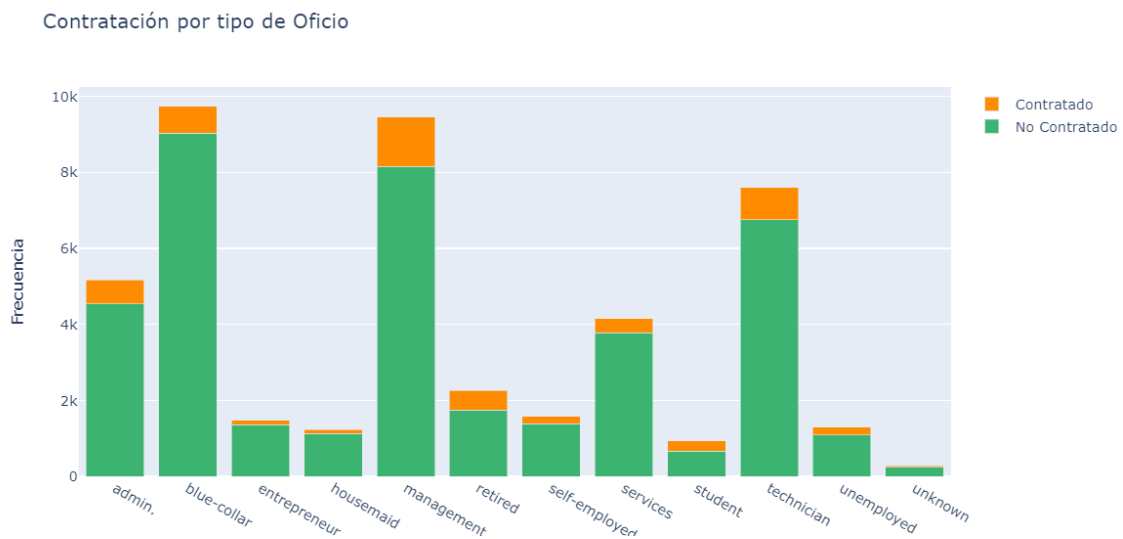
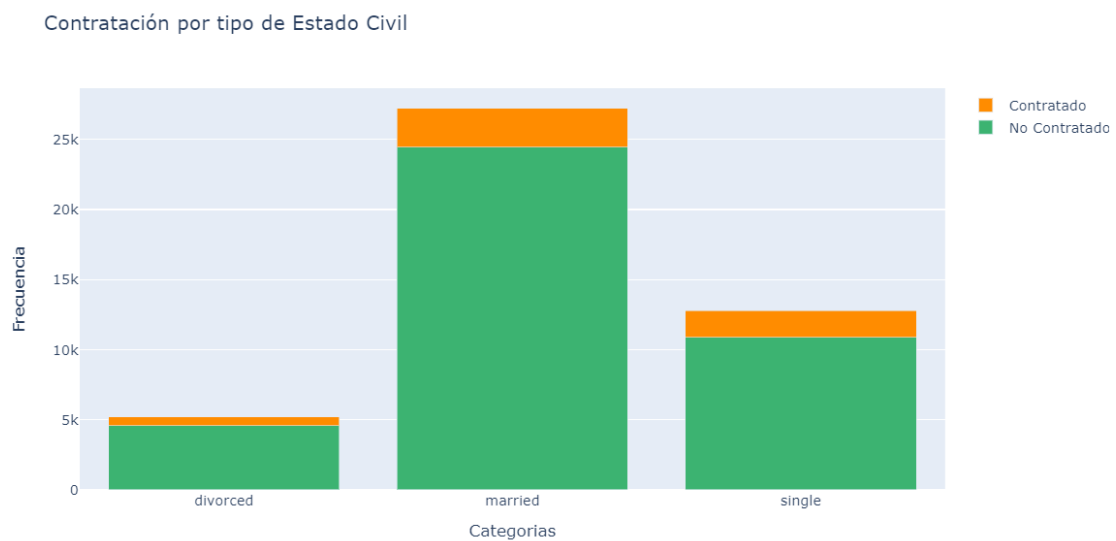


Análisis concentración de Clientes Contratados en función del tipo de variable

Inicialmente se comenzó realizando un análisis descriptivo de las variables explicativas en función de si un usuario fue contratado o no por la compañía.

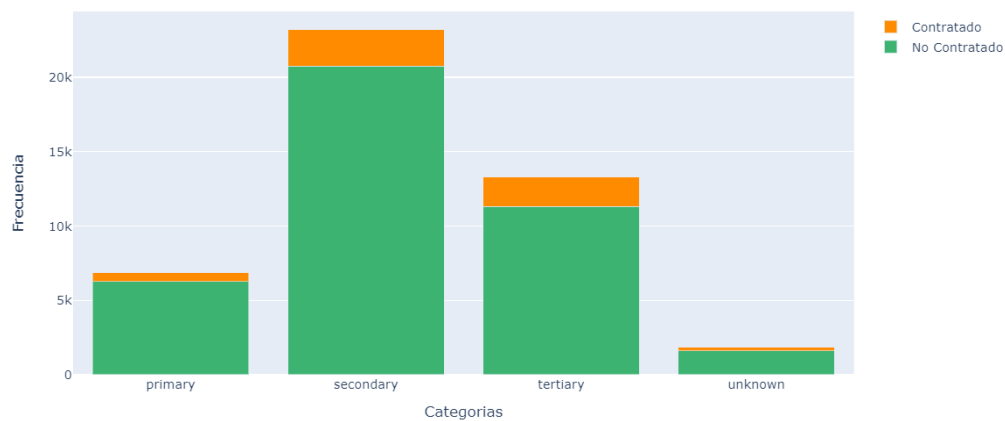


Como se puede observar, para la variable oficio, en la mayoría de los casos en la base de datos no se han contratado usuarios. No obstante, se puede observar que en los puestos directivos como en los puestos técnicos hay una mayor cantidad de contrataciones respecto a otras categorías.



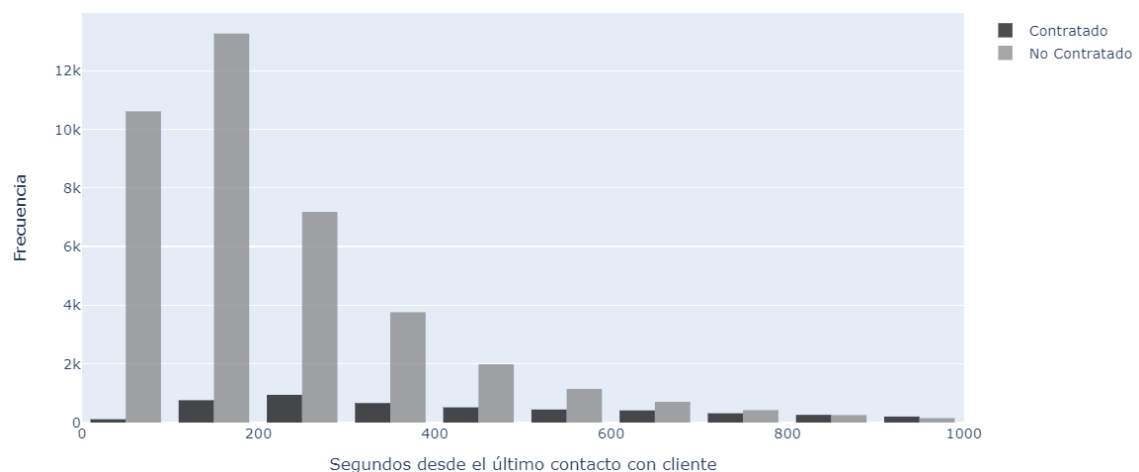
Siguiendo un análisis similar pero considerando la variable Estado Civil de los usuarios considerados en la base de datos, podemos observar que la mayoría de contrataciones se encuentran en aquellos usuarios casados. Las personas divorciadas tienen menos probabilidades de ser contratadas que otros estados civiles.

Contratación por tipo de Nivel De Estudios



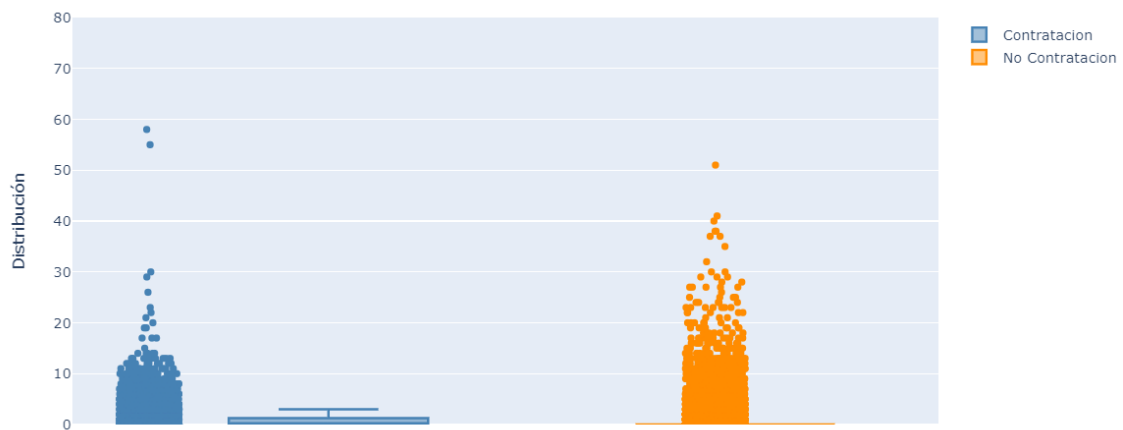
Por último, aquellas personas con estudios secundarios o terciarios tienen más probabilidades y son más propensas a ser contratadas respecto a otro tipo de nivel de estudios.

Distribucion de la Duracion en Función de la Contratación



En el histograma presentado se puede observar que independientemente de si el cliente ha sido contratado o no, el tiempo desde la última vez que un cliente fue contactado no varía. Adicionalmente, podemos observar en ambos casos que la mayoría de los contactos en la base de datos se produjeron en corto periodo de tiempo. Esperamos que esta variable no tenga mucho peso en el modelo predictivo del modelo dada la indiferencia entre las distribuciones, adicionalmente, esta gráfica alumbra que la insistencia en el cliente no es un factor diferencial para su contratación.

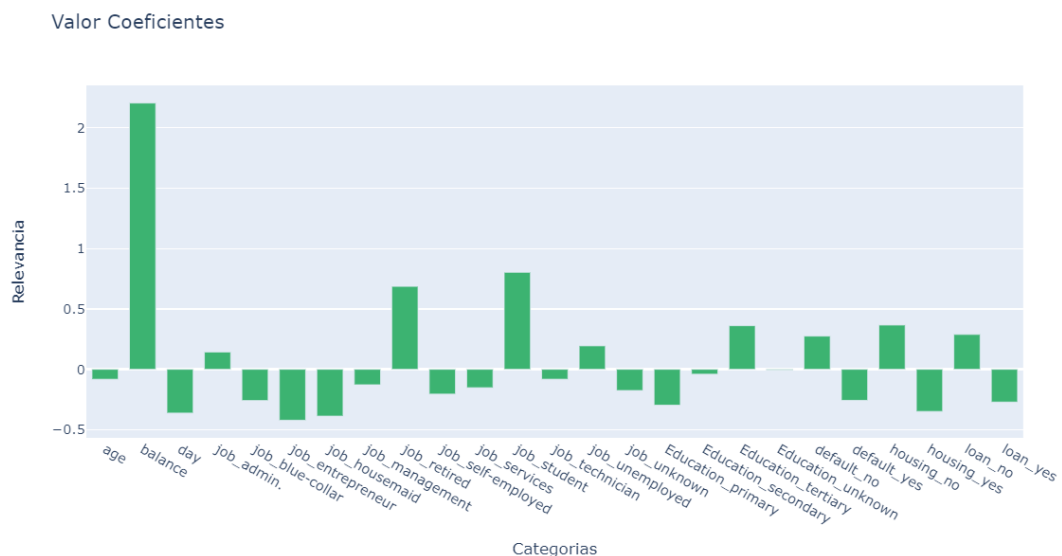
Contactos con el cliente en Función de la Contratación



A continuación, se presenta un diagrama de cajas con las distribuciones en función del número de contactos con el cliente. En este caso se puede observar que tampoco es diferencial los contactos con el cliente. Debido al amplio número de No contratados, vemos que ambas distribuciones prevalecen las mismas formas. Podemos concluir a partir de la gráfica que no hay un número de contactos óptimo para garantizar las contrataciones de los clientes.

Esperamos que aquellas variables con más cantidad de síes tengan un mayor coeficiente en el modelo de regresión presentado posteriormente. Adicionalmente, se puede observar como conclusión general que la mayoría de las muestras son NO, y que, en métricas agregadas, salvo en los patrones mostrados en los gráficos de barras, no hay variables diferenciales que nos indiquen potencialmente que estén muy correlacionadas con la salida.

Modelo predictivo



Tras normalizar los valores y construir variables dummy para predecir la variable explicativa (Si un usuario es contratado o no), se puede observar que tanto el balance como el tipo de oficio del usuario son las variables más determinantes para considerar si un usuario será contratado

o no. Variables como el estado civil o el tipo de educación, no han resultado ser tan determinantes como otras variables estudiadas.

Por último, con el fin de caracterizar el modelo y dar legitimidad a las conclusiones extraídas anteriormente. Podemos observar que la regresión lineal construida predice generalmente que un usuario no va a ser contratado. Esta tendencia seguramente se deba al gran número de muestras de No contrataciones observadas a lo largo del análisis. Por tanto, debe ser considerado como más fiable el análisis descriptivo de las variables más que el valor de los coeficientes del modelo de regresión logística debido a la poca capacidad predictiva del modelo.