instituto
de ingeniería
del conocimiento

20
21

N°
01

Revista de Casos de Estudio en HR Analytics

Journal of HR Analytics Case Studies

ISSN: 2792-3770 | Instituto de Ingeniería del Conocimiento

04 Estudio de la brecha salarial: un análisis de cómo impacta el género sobre el salario.

Gender Pay Gap: an analysis on how gender impacts salary.

ISSN: 2792-3770

La Revista de Casos de Estudio en HR Analytics

Journal of HR Analytics Case Studies

La ***Revista de Casos de Estudio en HR Analytics*** nace con la misión de facilitar el intercambio de conocimiento especializado entre profesionales y académicos en el ámbito de la **analítica de Recursos Humanos**, con el objetivo de mejorar la **efectividad de las organizaciones.** La entidad responsable de esta revista es la **Asociación para el Desarrollo de la Ingeniería del Conocimiento** (ADIC), siendo esta publicación on-line editada por el **Instituto de Ingeniería del Conocimiento** (IIC) con una periodicidad de un número anual.

Objetivo

La revista tiene como **objetivo** principal ser un vehículo para la reflexión y la difusión de las **buenas prácticas, últimos avances ylíneas de investigación** en el ámbito de la analítica aplicada para la toma de decisiones sobre la gestión del capital humano en las organizaciones.

La revista tiene un **carácter científico** y una **vocación divulgativa**, por ello propone artículos fundamentalmente de **carácter aplicado**. Con ellos se pretende que los profesionales de las organizaciones accedan a un conocimiento relevante acerca de cómo otras organizaciones desarrollan HRA. Y, también, acercar a los académicos el conocimiento respecto de cómo se desarrolla HRA en la práctica.

Alcance

El **enfoque de la Revista**, que pretende ser **multidisciplinar**, da cabida (entre otros) a manuscritos que: reflejen **casos prácticos** de aplicación del HRA en las organizaciones; que analicen, comparen y relacionen la utilidad de diferentes **técnicas y/o herramientas** para el abordaje de diferentes objetivos analíticos; que planteen y valoren la efectividad de diferentes **metodologías de trabajo** para el desarrollo de proyectos HRA; que ayuden a entender el **mapa de ruta** por el que transitar desde los niveles básicos del HRA hasta los niveles de excelencia; y que en general ayuden a entender cómo **mejorar la efectividad organizacional** a partir de la analítica de datos referidos a la fuerza de trabajo.

ISSN: 2792-3770

© ADIC - IIC

instituto de ingeniería del conocimiento

Equipo Editorial

La revista está editada por el Instituto de Ingeniería del Conocimiento y tiene los siguientes órganos de gobernanza.

Editor

David Aguado.
Instituto de Ingeniería del Conocimiento.

Editores Asociados

Jesús de Miguel.
Centro de Investigación para la Efectividad Organizacional, Universidad Autónoma de Madrid.

Antonio Delgado.
Universidad Autónoma de Madrid.

María Jesús Belizón.
University College Dublin.

Beatriz Lucía.
Instituto de Ingeniería del Conocimiento.

Delia Majarín.
Telefónica.

Sergio Raja.
Zurich Seguros.

Comité Editorial

Magdalena Nogueira.
Universidad Autónoma de Madrid.

Francisco Abad.
Universidad Autónoma de Madrid.

Carmen García.
Universidad Autónoma de Madrid.

José Manuel de Haro.
Universidad de Alicante.

William Ferrando Durán.
Universidad Javeriana.

Carolina Zúñiga.
Universidad Politécnica Salesiana del Ecuador.

José Carlos Andrés.
Viewnext.

Eduardo Páez.
Cepsa.

Pablo Haya.
Instituto de Ingeniería del Conocimiento.

Álvaro Barbero.
Instituto de Ingeniería del Conocimiento.

Sonia Rodríguez.
Instituto de Ingeniería del Conocimiento.

Celia Martínez.
Instituto de Ingeniería del Conocimiento.

Maite Sáez.
Observatorio de Recursos Humanos y Relaciones Laborales.

Diseño y Maquetación

Nuria Herranz González .
Instituto de Ingeniería del Conocimiento.

Andrés Muñoz Bachiller.
Instituto de Ingeniería del Conocimiento.

ISSN: 2792-3770

Índice de contenidos

00. La Revista de Casos de Estudio en HR Analytics al servicio de la comunidad profesional. 07

The Journal of HR Analytics Case Studies at the service of the professional community. As an Editorial in its first issue.

01. Aprendiendo a través de la práctica profesional del HRA: Introducción al primer número de la Revista de Casos de Estudio en HR Analytics. 13

Learning Through HRA Professional Practice: Introduction to the First Issue of the Journal of HR Analytics Case Studies.

02. Estudio de la Desigualdad de Género en una Empresa Constructora Española. 19

Study of Gender Inequality in a Spanish Building Company

03. Análisis de la Rotación de Personal para Mejorar el Proceso de Toma de Decisiones en una Empresa del Sector TIC: el caso de GFT IT Consulting, S.L.U. 41

Analysing staff turnover to improve the decision-making process in an ICT sector company: the case of GFT IT Consulting, S.L.U

04.Estudio de la brecha salarial: un análisis de cómo impacta el género sobre el salario 69

Gender Pay Gap: an analysis on how gender impacts salary

05. La elaboración del Informe de Estado de Información No Financiera Paso a Paso: La Experiencia de Grupo Santalucía, en cuestiones relativas al personal. 85

Preparation of the Non-Financial Information Status Report Step by Step: The Experience of Grupo Santalucía

06. Medición del impacto en el negocio al poner en valor a las personas: el caso MAPFRE Ser para el Hacer. Una alternativa analítica sin el uso de grupos pre y post. Primera aproximación 105

Impact Measurement on the business by adding value in the people: the MAPFRE Be for Doing case. An analytical alternative without the use of pre and post in. First approach

07. Explorando los determinantes del compromiso organizacional a través de las competencias y las motivaciones individuales: estudio HRA del compromiso corporativo. 123

Exploring the determinants of organizational engagement through skills and individual motivations: HRA Study of Corporate Engagement.

08. Desde la Academia: 8 *Insights* para construir un HRA efectivo. 143

From the Academy: 8 Insights to Build an Effective HRA.

ISSN: 2792-3770

© ADIC - IIC

ISSN: 2792-3770

04

Estudio de la brecha salarial: un análisis de cómo impacta el género sobre el salario

Gender Pay Gap: an analysis on how gender impacts salary

ISSN: 279

Nathan A. Consuegra Rengifo
Telefónica

Correo electrónico:
nathan.consuegrarengifo@telefonica.com

Linkedin:
https://www.linkedin.com/in/nathan-consuegra/

Delia Majarín Zacarías
Telefónica

Correo electrónico:
delia.majarinzacarias@telefonica.com

Linkedin:
https://www.linkedin.com/in/deliamajarin/

© ADIC - IIC

ISSN: 2792-3770

Palabras clave:
Brecha salarial, compensación, People Analytics, HR Analytics, Regresión.

Resumen

Reportar la brecha salarial entre hombres y mujeres se les solicita a todas las empresas con sede en España desde finales de 2018 lo que supone un reto para las empresas que demanda una recogida de información y un análisis detallado. En este caso de uso de brecha salarial se expone un análisis de la brecha salarial ajustada mediante la que, usando técnicas de regresión múltiple, se explica el salario a través de diferentes factores relevantes para la compensación, incluyendo el género, para así atribuir cuál es el efecto del género en la definición del salario de un empleado.

Abstract

Reporting gender pay gap is mandatory for all companies based in Spain since the end of 2018. It represents a challenge for companies that requires information collection and detailed analysis. In our case of use of gender pay gap, an analysis of the adjusted salary gap is presented through which, using multiple regression techniques, salary is explained through different factors relevant to compensation, including gender, in order to attribute which is the effect of gender on the definition of an employee's salary.

Keywords:
gender pay gap, compensation, People Analytics, HR Analytics, regression model

1. Introducción

Telefónica es una compañía del sector de las Telecomunicaciones cuyo objetivo es facilitar la comunicación entre las personas, proporcionándoles la tecnología más segura y de vanguardia, para que vivan mejor y consigan lo que se propongan. Telefónica opera en 12 países y tiene presencia en 24 con un promedio de más de 110.000 empleados.

Su *headquarter* está en Madrid, así como la mayor parte de su equipo corporativo de Recursos Humanos en el que se definen los planes estratégicos del área, así como se lideran las acciones y proyectos que impulsan la transformación de esta, siempre con el objetivo de prestar un servicio excelente al negocio, y manteniéndose en la vanguardia de cualquier nueva tendencia que surja en el ámbito de Recursos Humanos que pueda ayudar a cumplir este objetivo.

Una de las palancas estratégicas de transformación en el área de Recursos Humanos son los datos y la tecnología. No es ajeno a ningún departamento de Recursos Humanos la importancia del *reporting*, el seguimiento de KPIs estratégicos o la elaboración de cuadros de mando para asegurar que las acciones que se ejecutan en Recursos Humanos están alineadas con la estrategia del negocio y la compañía. Sin embargo, en 2017 esta práctica que ya venía siendo habitual, fue impulsada por una nueva tendencia que comenzó a surgir en las organizaciones y supuso que en Telefónica se pusiese en marcha una estrategia de People Analytics.

Muchos autores han descrito esta disciplina pero quizá una definición que se aproxima mejor a la forma de entender y ejecutar People Analytics en Telefónica sea la de Marler & Boudreau (2017) en la que definen HR Analytics (término más usando en el ámbito académico) como "*HR practice enabled by information technology that uses descriptive, visual, and statistical analyses of data related to HR processes, human capital, organisational performance, and external economic benchmarks to establish business impact and enable data-driven decision-making*". En cualquier caso, para Telefónica, el foco está en poder tomar decisiones basadas en los datos, aportar insights de valor al negocio e, incluso, crear productos analíticos que mejoren la experiencia de empleado. Para conseguir este objetivo, en los últimos años diversas iniciativas se han puesto en marcha, desde la creación de un *data lake* de datos de empleados llamado #OnHR hasta el llevar a cabo diferentes casos de uso según las prioridades del momento de Recursos Humanos y del negocio. El caso de uso que se expone a continuación es fruto de este esfuerzo por tener a disposición los datos clave de los empleados, con la calidad suficiente para poder explotarlos y ayudarnos a tomar mejores decisiones, basadas en evidencias.

1.1. Contexto Legal

Reportar la brecha salarial entre hombres y mujeres se les solicita a todas las empresas con sede en España en la **ley 11/2018 del 28 de diciembre,** la cual modifica el Código de Comercio en materia de información no financiera y diversidad. En ella, se recogen los siguientes principios:

- Ambiente diverso y cambiante. Es importante **garantizar la igualdad de oportunidades** y el mismo tratamiento **sin ningún sesgo por género.**
- **La Unión Europea promueve la igualdad entre hombres y mujeres,** y refuerza la igualdad como uno de los principales valores. Esto se refleja en las Cartas de los Derechos Fundamentales y en el Tratado de la Unión Europea.

Para cumplir con lo establecido en la ley y poder aportar una métrica que pueda representar la igualdad entre hombres y mujeres, la compañía ha optado por seguir el ejemplo de Reino Unido debido a su mayor avance en cuanto a legislaciones relacionadas con igualdad de género. El criterio

ISSN: 2792-3770

© ADIC - IIC

establecido en este país durante el año 2017 define el concepto de **brecha salarial bruta como "la diferencia entre la remuneración de los hombres y mujeres** expresada de forma relativa a la remuneración del hombre". Sin embargo, la definición de brecha salarial bruta no toma en consideración otros factores que tienen impacto en la remuneración percibida por los empleados de la compañía, como lo pueden ser la categoría profesional, la antigüedad o la edad. Por lo tanto, y a modo de explicación del resultado de la **brecha salarial bruta,** en la compañía se ha optado por incluir el cálculo de la **brecha salarial ajustada**, cuyo fin "consiste en explicar la diferencia salarial entre hombres y mujeres e indicar qué porcentaje de brecha no se puede explicar con la inclusión de nuevos factores".

1.2. Contexto Analítico

Durante el mes de diciembre de 2018, la compañía contrató a una consultora como asesor en el estudio de la brecha salarial entre hombres y mujeres. La colaboración de la empresa ayudó a la organización a sentar las bases en cuanto al tipo de información que se debía recopilar para el estudio y en la definición del cálculo de las brechas brutas y ajustadas.

En el caso de la brecha salarial ajustada, tanto la consultora contratada como un informe publicado por Glassdoor en el año 2017 recomiendan el uso de **modelos de regresión múltiple** para explicar la brecha salarial bruta de la compañía. Este tipo de modelo permite explicar el efecto que un conjunto de factores, denominados *variables independientes,* puede tener en otro factor denominado *variable dependiente*. En el caso del cálculo de la brecha salarial ajustada, la variable dependiente es el total de los conceptos salariales percibidos por un empleado, mientras que las variables independientes son todas aquellas que puedan explicar el salario como lo son la categoría profesional, la antigüedad, la edad, etc... El género siempre se debe incluir como variable independiente para poder evaluar su impacto en el salario de un empleado.

1.3. Objetivo del Estudio

Indicar con total transparencia cual es la diferencia salarial entre hombres y mujeres a través del cálculo de la brecha salarial bruta tomando en consideración todos los conceptos retributivos percibidos por los empleados de la organización.

1. Explicar el porqué de la diferencia salarial entre hombres y mujeres a través del cálculo de la brecha salarial ajustada e indicar, en el caso de que corresponda, cuál es el efecto del género en la definición del salario de un empleado.
2. Fijar las bases para el cálculo del **Equal Pay.** Para esto, el cálculo de la brecha ajustada se ha definido como un proceso iterativo en el cual se pretende incluir nuevos factores año tras año para poder aumentar el nivel de explicación y llegar a medir el Equal Pay, el cual indica si hombres y mujeres en el mismo rol y nivel de desempeño reciben el mismo pago.

ISSN: 2792-3770

2. Método

Desde que la ley se publicó en diciembre de 2018, el ejercicio se realiza anualmente durante el mes de enero del año siguiente y con información recogida con fecha de corte a 31 de diciembre del año en curso. Con la información recopilada se realizan 2 cálculos:

- **Brecha Salarial "Bruta" entre Hombre y Mujeres:** consiste en calcular la diferencia porcentual entre la remuneración media percibida por los hombres con respecto a las mujeres. El cálculo se realiza con la siguiente fórmula:

$$\overline{Remuneracion_{Hombre}} = \text{Promedio de la Remuneración percibida por los Hombres}$$
$$\overline{Remuneracion_{Mujer}} = \text{Promedio de la Remuneración percibida por las Mujeres}$$

$$Brecha\ Bruta = \frac{\overline{Remuneracion_{Hombre}} - \overline{Remuneracion_{Mujer}}}{\overline{Remuneracion_{Hombre}}}$$

- **Brecha Salarial "Ajustada" entre Hombre y Mujeres:** consiste en explicar cómo se define la remuneración total de un empleado en base a diferentes factores, incluido el género. Para esto, se genera un *modelo de regresión múltiple* (Diez 2019) donde la *variable dependiente* es la **remuneración total** de un empleado y las variables independientes son **factores demográficos** relacionados con el puesto laboral del trabajador. El modelo se define con la siguiente ecuación:

$$log(Remuneración\ Total) = \beta_0 + \beta_1 * Género + \beta_2 X_2 + \cdots + \beta_i X_i + \epsilon$$

Donde el efecto del género en el salario de un empleado se determina a través del coeficiente β_1, el cual será reducido debido a la presencia de los coeficientes β_i de los diferentes factores X_i que se incluyan en el cálculo. β_0 es el c*oeficiente constante* o *intercepción* que indica cómo se comporta el salario según los valores por defecto de cada variable X_i y ϵ es el error. El motivo de la transformación logarítmica del salario se explica en el apartado 2.3.

Dado que el fin principal de este cálculo es *explicar la brecha salarial bruta* de la compañía, en el futuro siempre se podrá añadir nuevos factores X_i que ayuden a explicar la remuneración percibida por los empleados.

2.1. Población

Al ser la compañía donde se realiza el estudio una multinacional con sede en España, la ley exige que se realice el estudio con la población de empleados pertenecientes a las diferentes empresas repartidas en 5 grandes regiones del planeta: España, Alemania, Reino Unido, Brasil e Hispanoamérica. Por lo tanto, en total se tiene una población de unos 114.000 empleados aproximadamente donde sólo se toman en cuenta *empleados activos o con baja remunerada* con fecha de corte a 31 de diciembre del año para el que realice el ejercicio.

ISSN: 2792-3770

© ADIC - IIC

2.2. Sistemas, Herramientas, Datos y Modelos

2.2.1. Sistemas y Herramientas

La información de los empleados es enviada por los responsables de planificación y compensación de cada una de las operadoras que conforman la compañía. Para esto, la empresa dispone de una **aplicación web** propia donde los responsables pueden cargar la información a través de un fichero Excel.

Una vez recopilada esta información en el sistema, se puede exportar en formato *csv* y utilizar un script en **Python v3.5** para realizar el cálculo de la brecha salarial ajustada. El programa utiliza las siguientes librerías para poder ejecutar el modelo:

- **Pandas v1.1.5:** permite leer y manipular el contenido de ficheros con formato csv, excel u otros.
- **Statsmodels v0.12.1:** módulo que provee diferentes clases y funciones para la estimación de diferentes modelos estadísticos.

2.2.2. Datos y Modelos

Durante el momento del ejercicio se ejecutan 2 procesos distintos para recopilar la información necesaria para el estudio:

- **Planificación (PLN):** este proceso se realiza previo al lanzamiento de la ley sobre la brecha salarial y consiste en recopilar la información demográfica de todos los empleados de la empresa.
- **Recopilación de la Información Salarial (GPG):** se realiza desde el momento de publicación de la ley y consiste en recopilar todos los conceptos salariales percibidos por todos los empleados de la compañía durante desde el *1 de enero hasta el 31 de diciembre del año en que se realiza el ejercicio*. La recopilación de estos conceptos se realiza en moneda local, pero para los ejercicios se transforman o a euros según el tipo de cambio correspondiente a la moneda origen.

ISSN: 2792-3770

El listado de las variables recopiladas a través de estos procesos, así como una breve explicación de estas, se puede observar en la tabla 1.

Tabla 1. Listado de las variables recopiladas a través de estos procesos de Planificación (PLN) y Recopilación de la Información Salarial (GPG)

PROCESO	VARIABLE	DESCRIPCIÓN
PLN	COUNTRY	País donde trabaja el empleado
	GENDER	Género del empleado: M o F
	AGE	Edad del empleado (años)
	SENIORITY	Antigüedad en la empresa (años)
	FUNCTIONAL_AREA	Área de trabajo
	LOCAL_PROFESSIONAL_CATEGORY	Categoría profesional de la empresa específica
	GLOBAL_PROFESSIONAL _CATEGORY	Categoría profesional de la compañía
	LEGAL_ENTITY	Jurídica a la que pertenece el empleado
	REGION	Región dentro del país donde trabaja el empleado.
	COLLECTIVE_AGREEMENT	Acuerdo laboral que posea el empleado
	EMPLOYEE_TYPE	Tipo de jornada: parcial o completa
	EMPLOYEE_CLASS	Tipo de empleado: practicante, expatriado, temporal, etc..
	EMPLOYEE_STATUS	Estado del empleado: activo, baja remunerada o baja no remunerada
GPG	ANNUAL_BASE_SALARY_EUR	Salario base anualizado
	REAL_PAID_BONUS_EUR	Bono cobrado
	REAL_PAID_COMMISSIONS_EUR	Comisiones cobradas
	PSP_TFSP_EUR	PSP/TFSP liquidados
	RSP_OTHERS_ILPs_EUR	RSP/Otros ILPs liquidados
	MEAL_TICKETS_EUR	Cheques restaurante
	HEALTH_INSURANCE_EUR	Seguro médico
	LIFE_INSURANCE_EUR	Seguro de vida
	ACCIDENT_INSURANCE_EUR	Seguro contra accidentes
	CAR_ALLOWANCE_EUR	Prestaciones por coche
	ATAM_EUR	ATAM
	ALBENTURE_EUR	Albenture
	TRANSPORT_PLUS_EUR	Prestaciones por transporte
	PCF_STOCK_EUR	Bolsa PCF
	AVAILABILITY_PLUS_EUR	Plus Disponibilidad
	MOVISTAR_SHARED_PAYMENT_EUR	Pago compartido Movistar
	BENEFIT_PER_CHILD_EUR	Ayuda por hijo
	STAFF_RELATED_COSTS_EUR	Formación, lugar de trabajo, etc…
	OTHER_BENEFITS_EUR	Suma de otros beneficios locales
	OTHER_CONCEPTS_EUR	Suma de otros conceptos locales

ISSN: 2792-3770

© ADIC - IIC

Dado que los empleados de la compañía perciben una gran variedad de conceptos salariales, para el cálculo se genera la variable **Remuneración Total**, cuyo valor será la suma de todos los conceptos salariales recopilados en el proceso de GPG. Se utilizará tanto el cálculo de la brecha salarial bruta como en el cálculo de la brecha salarial ajustada, siendo el factor dependiente en esta última.

El modelo definido en el programa se entrena para cada uno de los países donde se ubican las principales operaciones de la compañía: España, Alemania, Reino Unido, Brasil e Hispanoamérica. Los factores explicativos incluidos dentro del modelo de cada país son los siguientes:

- **AGE:** está estrechamente relacionada con la categoría profesional del empleado, ya que, de forma general, a mayor edad mayor rango y por tanto mayor salario.
- **SENIORITY:** En el caso de entidades jurídicas, los empleados con mayor tiempo dentro de la compañía suelen recibir un mayor salario que el resto.
- **LOCAL_PROFESSIONAL_CATEGORY:** los salarios son más elevados cuando la categoría profesional es alta debido al incremento de responsabilidades que una subida de categoría conlleva. Por lo tanto, los directores, gerentes y jefes perciben una mayor remuneración que el resto de los empleados.
- **FUNCTIONAL_AREA:** el tipo de trabajo desempeñado por el empleado también un impacto en su salario, ya que no gana lo mismo una persona que trabaja dentro del departamento de redes que una persona que trabaja dentro del departamento de finanzas. Existen varias encuestas de mercados que indican como el área funcional impacta la remuneración de los empleados como es la de Willis Towers Watson.
- **LEGAL_ENTITY (A partir del 2019):** la compañía se compone de diferentes entidades legales o jurídicas y cada una de ellas ofrece un tipo de contrato diferente, por lo que la remuneración de un empleado varía según la jurídica a la que pertenezca.

En el caso de la operación alemana, existe otros factores que ayudan a explicar la brecha salarial entre hombre y mujeres, por lo que se incluyen dentro del cálculo. Como estas variables sólo estaban disponibles para las jurídicas alemanas, no se han podido incluir en los modeles del resto de países para los ejercicios del 2018 y 2019. Estos factores son los siguientes:

- **REGION:** la remuneración varía según la ciudad donde trabaje el empleado.
- **COLLECTIVE_AGREEMENT:** de forma similar a la entidad legal, el acuerdo laboral define los beneficios que un empleado recibirá.
- **EMPLOYEE_TYPE:** determina si un empleado trabaja a tiempo completo o parcial. Trabajar a tiempo parcial suele implicar una reducción salarial.

2.3. Transformaciones e Interpretación

Los modelos de regresión requieren que las muestras sigan una distribución normal. A esto se le conoce como *asunción de normalidad.* Debido a esto, es una práctica común aplicar

ISSN: 2792-3770

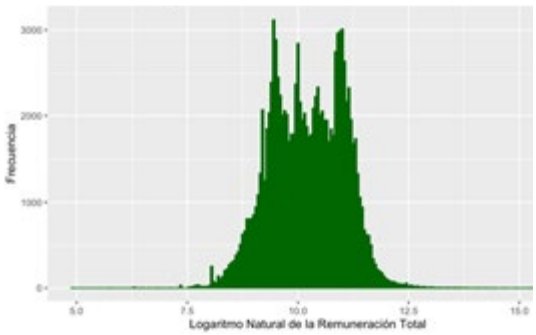
transformaciones tanto en las variables dependientes como independientes.

En el caso de la remuneración total de los empleados, se puede observar como su distribución presenta una *asimetría positiva* bastante pronunciada. Esto se debe a que los salarios altos se encuentran en altos cargos donde hay menos personas, mientras que los salarios más bajos se encuentran en la mayor parte de la población.

Figura 1.

Este tipo de distribución no es la más adecuada para un modelo de regresión. Por lo tanto, se optó por aplicar una *transformación logarítmica natural* a la remuneración total del empleado. En la siguiente gráfica se puede observar como esta transformación permite acercar la distribución de los salarios a una distribución normal, logrando con esto eliminar la asimetría positiva observada en la distribución original.

Figura 2.



Sin embargo, aplicar una transformación logarítmica no sólo tiene el beneficio de generar una distribución normal en la remuneración total, sino que además tiene un impacto en la *interpretación de los resultados.* Gracias a esta transformación, y siempre que los coeficientes sean próximos a 0, el efecto de las variables independientes se puede interpretar como cambios porcentuales en la remuneración total de los empleados. Es decir:

$$Si\ \beta \approx 0 \rightarrow e^{\beta} \approx 1 + \beta$$

Como la finalidad del estudio es indicar la diferencia porcentual entre la remuneración media percibida por los hombres con respecto a la percibida por las mujeres, el efecto del *género* se podrá expresar como porcentaje aplicando el siguiente calculo:

$$100 * (e^{\beta} - 1)$$

Dado que el género es una variable categórica con sólo 2 valores posibles (Masculino o Femenino), la variable se *dummifica*, es decir, se transforma en valores numéricos donde el género femenino se codifica con 0 mientras que el género masculino con 1. De esta forma, la parte de la ecuación del modelo que hace referencia al género $(\beta * Género)$ tiene la siguiente interpretación:

- **Si género femenino:** el producto entre el coeficiente y la variable da como resultado 0, por lo que el efecto de ser mujer se recoge en el *coeficiente constante* o *intercepción* (β_0) del modelo.
- **Si género masculino:** el producto entre el coeficiente y la variable es diferente de 0 e indica la diferencia porcentual entre la remuneración total percibida por los hombres con respecto a la percibida por las mujeres. Si el coeficiente es positivo, la interpretación es que el hombre percibe una remuneración mayor a la de la mujer. Si es negativo, la interpretación es que el hombre percibe una remuneración menor.

ISSN: 2792-3770

© ADIC - IIC

De igual forma, la dummificación aplicada a la variable *Género* se aplica a todas las variables categóricas presentes en el modelo. En concreto, se aplica en las variables: Categoría Profesional Local, Área Funcional, Entidad Legal, Región, Acuerdo Laboral y Tipo de Empleado. Esta acción la realiza la librería de statsmodels automáticamente cuando detecta que el tipo de variable es categórica y no numérica.

2.4. Evaluación

Para que un modelo de regresión sea considerado satisfactorio, éste debe cumplir con una serie de asunciones que se pueden observar en la Tabla 2.

Tabla 2. Supuestos para la valoración de la calidad de los modelos de regresión

ASunción	Descripción
Independencia de los Residuos	Los residuos son la diferencia entre el valor real de la observación y los *valores ajustados*, es decir; los valores devueltos por el modelo una vez aplicado el efecto de los coeficientes. Se considera que un modelo de regresión satisface esta asunción cuando los residuos no forman ningún tipo de *patrón*. Si esto ocurre, entonces sería síntoma de que el modelo no ha capturado algún comportamiento de la población y por tanto no es explicativo de esta. Tests estadísticos como Durbin-Watson y Jarque-Bera pueden verificar si un modelo satisface esta asunción.
Normalidad de los Residuos	Los residuos deben seguir una distribución normal. De no hacerlo, sería otro indicativo de que existe algún patrón que no ha sido detectado por el modelo. El test estadístico Jarque-Bera puede ser utilizado para confirmar esta asunción.
Homocedasticidad de los Residuos	Consiste en verificar si los residuos presentan una varianza constante para cada nivel de los valores ajustados. Cuando esto no se cumple, la varianza de los coeficientes aumenta lo cual implica que el modelo sera menos confiable. Los tests estadísticos Breusch-Pagan y Goldfeld-Quandt son útiles para verificar esta asunción.
Significancia estadística de los factores	Para cada una de las variables independientes se debe evaluar cual es su *significancia estadística o p-valor.* Cuando este valor se encuentra por debajo del 0.05, se considera que la variable es significativa y por tanto relevante para la explicación de la variable dependiente.

Otras medidas que se toman en consideración para evaluar el modelo son el R^2 y el $R^2\ ajustado$. Estas métricas permiten evaluar cual es el grado de efectividad con el cual las variables independientes del modelo pueden explicar la varianza de la variable dependiente. Su valor está ubicado en el intervalo [0,1], siendo deseable obtener un valor próximo a 1.

ISSN: 2792-3770

3. Resultados

Desde el momento en que se publicó la ley, el ejercicio se ha realizado 2 veces. El primero, denominado ejercicio A, en enero de 2019 con datos de empleados a 31 de diciembre de 2018 y el segundo, denominado ejercicio B, en enero de 2020 con datos de empleados a 31 de diciembre de 2019. Para explicar los resultados, tomaremos en cuenta el caso de 2 países que componen la compañía: España y Alemania

La tabla 3 proporciona un resumen de los resultados de las brechas brutas y ajustadas de ambos ejercicios y países. La significación estadística de las variables para todos los modelos y ejercicios de ambos países indica que todos los factores incluidos en los modelos de la brecha ajustada son *significativos.*

Tabla 3. Resumen de Resultados de la Brecha Salarial Bruta y Ajustada en España y Alemania

PAÍS	Ejercicio A (2018)		Ejercicio B (2019)	
	Brecha		Brecha	
	Bruta	Ajustada	Bruta	Ajustada
España	15.18%	4.50%	14.88%	3.41%
Alemania	24.63%	4.60%	22.75%	4.40%

Para el ejercicio A de ambos países, y debido a que la ley se publicó en diciembre de 2018, se contó con poco margen para recopilar y analizar la información demográfica y salarial de los empleados. Por lo tanto, sólo fue posible utilizar los siguientes factores, denominados factores base, para el cálculo de la brecha ajustada:

- **AGE**
- **SENIORITY**
- **FUNCTIONAL_AREA**
- **LOCAL_PROFESSIONAL_CATEGORY**

3.1. Resultados de España

Gracias a los factores base, el resultado del modelo para España indica un coeficiente para el género de $\beta_{Género} = 0.0440$. Aplicando la transformación $100 * (e^{\beta_{Género}} - 1) = 4.4498$, el modelo indica que para España país existe una brecha ajustada del 4.50% en el 2018. Dado que la brecha bruta del país de este año es del 15.18%, la conclusión de este análisis es que los factores de *edad, antigüedad, área funcional y categoría profesional* permiten explicar aproximadamente un 11% de la brecha bruta del país.

En cuanto al ejercicio B, no sólo se han tomado en cuenta los factores base, sino que también se incluyó el concepto de *LEGAL_ENTITY o Entidad Legal.* Con la incorporación

ISSN: 2792-3770

© ADIC - IIC

de este nuevo factor, el resultado del modelo para España indica un coeficiente de $\beta_{Género} = 0.0335$. La interpretación de este coeficiente en modo porcentual es de $100 * \left(e^{\beta_{Género}} - 1\right) = 3.4067$, lo cual implica un 3.41% de brecha ajustada en el 2019. El resultado de la brecha bruta para este ejercicio es del 14.88%, por lo que la inclusión del nuevo factor incrementa la capacidad de explicación de la brecha bruta hasta en un 0.50% aproximadamente con respecto al 11% del ejercicio anterior.

Los resultados de las pruebas estadísticas para ambos ejercicios se encuentran recogidos en la tabla 4, donde el resultado será el p-valor correspondiente a la prueba exceptuando la medida $\mathbf{R^2\ ajustado}$ y la prueba de correlación de Durbin-Watson:

3.2. Resultados de Alemania

Aplicando los factores base en el ejercicio A, el resultado del modelo para Alemania indica un coeficiente para el género de $\beta_{Género} = 0.0865$, lo cual implica una interpretación porcentual del $100 * \left(e^{\beta_{Género}} - 1\right) = 9{,}0351$. Es decir, un 9.00% de brecha ajustada en el modelo del 2018. La brecha bruta correspondiente a este país y ejercicio es del 24.63%, por lo que el modelo logra explicar aproximadamente un 15,30% de la diferencia salarial media entre hombres y mujeres de este país.

Los números indicados son el resultado obtenido por el equipo global de la compañía para el ejercicio A. Sin embargo, para el cálculo de la brecha ajustada, la operadora del país condujo un análisis interno donde no sólo se tomaron en cuenta los factores base, sino que además se incluyeron otros, denominados factores extra, que lograron reducir el efecto del género a un 4.60%. Estos son:

- **REGION**
- **COLLECTIVE_AGREEMENT**
- **EMPLOYEE_TYPE**

En cuanto al ejercicio B, el equipo global no sólo tomó en cuenta los factores base, sino que además incluyó los factores extra indicados por la operadora alemana durante el ejercicio A y el concepto de *LEGAL_ENTITY o Entidad Legal.* Aplicando e$\beta_{Género} = 0.0430$ obtiene un coeficiente de
, lo cual implica un efecto ajustado del género de $100 * \left(e^{\beta_{Género}} - 1\right) = 4.4393$. Es decir, un 4.40% de brecha ajustada para el modelo del 2019. La brecha bruta correspondiente a este país y ejercicio es del 22.75%, por lo que el modelo logra explicar aproximadamente un 18,30% de la diferencia salarial media entre hombres y mujeres de este país. Esto indica que la inclusión de los factores extra y el nuevo factor de entidad legal incrementa la capacidad de explicación de la brecha bruta hasta en un 3% aproximadamente con respecto al 15,30% del ejercicio anterior. Los resultados de las pruebas estadísticas para ambos ejercicios se encuentran recogidos en la tabla 4:

Tabla 4. Índices de Calidad de los Modelos de Regresión realizados para España y Alemania.

	R^2 Ajustado	Durbin-Watson	Jarque-Bera	Breusch-Pagan
Ejercicio A (2018) España	0.84	1.54	p-valor = 0.0	p-valor = 0.0
Ejercicio B (2019) España	0.89	1.77	p-valor = 0.0	p-valor = 0.0
Ejercicio A (2018) Alemania *	0.78	1.31	p-valor = 0.0	p-valor = 0.0
Ejercicio B (2019) Alemania	0.85	1.69	p-valor = 0.0	p-valor = 0.0

Nota: * Dado que el análisis realizado por la operadora de Alemania durante el ejercicio A no lo condujo el equipo global, no se tienen los resultados referentes ni a las pruebas estadísticas ni a la tabla de regresión, por lo que en esta Tabla sólo se recogen los obtenidos con los factores base.

ISSN: 2792-3770

3.3. Interpretación de los Resultados Estadísticos

Observando los resultados estadísticos para cada modelo, se observa que en términos de R^2 *ajustado* el menor de ellos es un 0.78, lo cual nos confirma que el modelo logra explicar una gran parte de la varianza de los datos y la inclusión de nuevos factores ha contribuido a este resultado. También se observan buenos resultados en el estadístico de Durbin-Watson dado que la mayoría de los valores se encuentran en el rango [1.5, 2.5], lo cual nos indica que los modelos, exceptuando el de Alemania para el ejercicio A, satisfacen esta asunción. La razón por la cual este modelo es el peor seguramente se deba a 2 razones:

- Para el ejercicio A, el modelo de Alemania no incluye los factores de región, acuerdo laboral y tipo de empleado. Esto se debe a que para cada región y acuerdo un empleado con las mismas condiciones y nivel de experiencia percibirá una remuneración total diferente. Además, el tipo de empleado indica si trabaja a tiempo completo o parcial. En el caso del último, implica una reducción en la remuneración del trabajador y es una situación donde el 77% de los empleados son del género femenino.
- En el desglose del factor de categoría profesional se ven detallados los diferentes puestos en altas categorías profesionales como gerentes y directores. Sin embargo, sólo la categoría *Employee* representa el 87% de los empleados en esta operadora, por lo cual el modelo no será capaz de diferenciar los diferentes tipos de niveles profesionales que existan dentro de ese conjunto. España no tiene este problema, dado que cuenta con una lista de 66 categorías profesionales diferentes donde el nivel con mayor cantidad de población representa un 21% del país.

Otro detalle para resaltar es que la mejora entre los ejercicios A y B se debe a la inclusión del factor de Entidad Legal en el modelo de regresión. Esto se debe a que cada jurídica dentro de la empresa ofrece beneficios salariales diferentes, por lo que incluir esta variable es imprescindible para explicar la distribución salarial de los empleados de la compañía.

3.4. Limitaciones del modelo

En cuanto a las pruebas realizadas para verificar las asunciones de Normalidad y Homocedasticidad de los residuos, estos son test estadísticos que confirman si podemos o no rechazar la hipótesis nula, la cual es que los *residuos son normales* en el caso de Jarque-Bera y que la *varianza es constante* en el caso de Breusch-Pagan. Se rechazará la hipótesis nula cuando el resultado de estos Tests indique un p-valor < 0.05. En el caso de los modelos descritos, todos rechazan la hipótesis nula para ambas pruebas. Las causas de esto pueden ser:

- No se han recopilado todas las variables que explican el salario. Un ejemplo puede ser el *performance* (desempeño profesional) del empleado.
- Dado que la variable dependiente es la remuneración total del empleado, puede que algunos conceptos salariales que no todos los empleados perciban por igual estén perjudicando al modelo. Las comisiones son un ejemplo de esto.
- La calidad del dato puede no ser correcta en todos los casos.
- No se han aplicado transformaciones a las variables independientes que ayuden a corregir los problemas del modelo.
- No se han eliminado outliers del estudio debido a que la ley exige que se tome en cuenta a todos los empleados activos o con baja remunerada con fecha de corte a 31 de diciembre.

ISSN: 2792-3770

© ADIC - IIC

4. Discusión

Previo al estudio de brecha salarial de género y a la obligatoriedad legal de presentar este resultado, la compañía ya había puesto en marcha planes de acción para fomentar la diversidad dentro de la empresa. Estas iniciativas no surgen debido al estudio de brecha, sino que son objetivos planteados por la compañía previos al año 2018 y que se han visto reforzados gracias al estudio, pero siempre han sido independientes a éste. El detalle se puede encontrar tanto en el sitio web dedicado a esta iniciativa como en el informe integrado publicado anualmente por la empresa.

El aplicar un modelo de regresión dentro del área de recursos humanos para explicar el problema de la brecha salarial entre hombres y mujeres permite demostrar de forma *cuantitativa* el impacto que los factores utilizados en el estudio tienen sobre el salario. Además, el tener este conocimiento dentro del área de recursos humanos otorga independencia como empresa al no necesitar de un tercero para su estudio.

Aunque el modelo de regresión presenta indicadores que muestran su buen rendimiento, es esperable que el modelo evolucione de año en año según se incluyan nuevas variables relevantes para explicar el salario, siendo este un proceso iterativo en el que nuestra aspiración es llegar a medir el Equal Pay. Esto dependerá en gran medida de la calidad del dato y la homogeneidad de estos entre los diferentes países que componen Telefónica. Para afrontar este gran reto se han puesto en marcha planes de Gobierno y Auditoría de los datos ya que esta credibilidad es una palanca fundamental para poder avanzar, no solo en este caso de uso analítico sobre brecha salarial sino en **People Analytics** en general. Además, y con el fin de ser lo más específico y cercano a la realidad posible, sería interesante estudiar cómo se encuentran las brechas salariales de género en cada área funcional para determinar si es necesario aplicar algún tipo de corrección específico a cada caso.

5. Limitaciones del estudio

Este estudio solo pretende explicar el impacto que el género tiene en la remuneración total de los empleados cuando se toman en consideración otros factores. Por la tanto, lo resultados sólo estarán relacionados con el impacto de esta variable.

A pesar de que la información recopilada para el estudio es auditada por una tercera entidad, la base de datos sigue teniendo un margen de error mínimo, por lo que la información no es 100% correcta para todos los registros.

Debido a que sólo se recopilan un número determinado de variables en los procesos de planificación y de información salarial, no puede incluirse otro tipo de variables que ayuden a explicar el salario del empleado como es el caso del performance, años de formación, centro educativo, experiencia previa u otros.

En el estudio no se pueden eliminar *outliers* debido a que la ley exige que se tomen en cuenta a todos los empleados activos o con baja remunerada.

Por último, los modelos descritos en este documento sólo cumplen algunas de las asunciones que los modelos de regresión múltiple deberían satisfacer.

ISSN: 2792-3770

6. Futuras líneas de investigación

Para el ejercicio del 2020, no sólo se incluirá el *tipo de empleado* como factor en el modelo de Alemania, sino que además se incluirá en el resto de los países y se estudiará su impacto en el salario. También se pretende incluir una nueva variable dicotómica que determine si un empleado ha cobrado comisión o no.

En cuanto a la mejora del modelo, se estudiará aplicar transformaciones a las variables independientes para corregir los problemas relacionados a la normalidad y homocedasticidad de los residuos.

Actualmente la empresa está mapeando las *skills o habilidades* que cada empleado posee en la compañía debido a que hoy en día en el mercado la remuneración del empleado puede variar según las skills que posea. Por lo tanto, se pretende incluir estos factores dentro del modelo ajustado para explicar la remuneración del empleado.

Por último, y con la intención de cumplir con lo exigido por la ley, actualmente en el departamento de recursos humanos de la empresa se está trabajando en el proyecto de *Auditoría Retributiva*. Su fin consiste en tener un registro donde se recopile tanto la información salarial del empleado como aquellos factores que identifiquen *puestos de igual valor*, y dar explicaciones cuando en determinados puestos se obtengan brechas salariales brutas que superen el 25%.

ISSN: 2792-3770

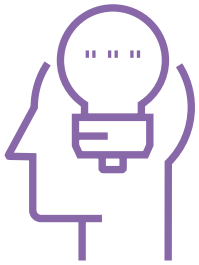
7. Conclusión

A través de técnicas utilizadas dentro del área de People Analytics, la compañía ha podido aplicar un modelo matemático con el fin de modelar el comportamiento de los empleados dentro de la empresa y con ello generar insights que explican un problema como es el de la brecha salarial de género. Con cada nuevo ejercicio realizado y factor añadido al modelo de regresión, la empresa no sólo ha podido comprobar **la diferencia entre la brecha salarial bruta y la real,** sino que además se acerca más al equal pay y con ello a poder demostrar que **el género no se toma en consideración para definir el salario de un empleado** y que, por tanto, no se tiene un problema de discriminación por género dentro de la compañía.

Los modelos descritos hasta ahora ayudan a explicar un gran porcentaje de la brecha salarial bruta y mejoran con cada nuevo avance, pero no son perfectos. Como se ha podido comprobar, aún no satisfacen todas las asunciones que un modelo de regresión debe tener. Esto se puede considerar como una buena noticia, ya que estos modelos incluyen al género como factor, lo cual nos dice que su presencia aún no implica un modelo 100% perfecto y se espera que cuando se tenga un modelo de estas características, el género no sea parte de la ecuación.

© ADIC - IIC

8. Referencias



Barr, C. D., Diez, D. M., Dorazio, L., & Cetinkaya-Rundel, M. (2019). *Advanced high school statistics.* OpenIntro.

Chamberlain, A. (2017). *How to Analyze Your Gender Pay Gap: An Employer's Guide.* Glassdoor Economic Research.

Marler, J. H., & Boudreau, J. W. (2017). An evidence-based review of HR Analytics. *The International Journal of Human Resource Management, 28(1)*, 3-26.

Willis Towers Watson (2017). Compensation Survey Report - U.S.

Telefónica(a). (2019). La diversidad es una fuente de talento https://www.telefonica.com/es/web/negocio-responsable/-/contenido-diversidad

Telefónica (b).(2019). Informe Integrado. https://www.telefonica.com/es/web/negocio-responsable/informe-2019

ISSN: 2792-3770

instituto de ingeniería del conocimiento

INSTITUTO DE INGENIERÍA DEL CONOCIMIENTO (IIC)

PIONEROS EN INTELIGENCIA ARTIFICIAL DESDE 1989

Somos un centro de I+D+i experto en **Big Data e Inteligencia Artificial**. El núcleo, experiencia y trayectoria del IIC gira en torno al análisis de datos.

Nuestra apuesta de valor se basa en el desarrollo de algoritmos y técnicas de análisis a medida, de modo que conformen soluciones tecnológicas altamente adaptadas a las necesidades de cada cliente.

Únete a un equipo joven y dinámico, formado por más de 150 profesionales especializados en tecnologías de vanguardia.
Estamos ubicados en la Universidad Autónoma de Madrid (UAM). Nos nutrimos del mejor talento universitario y somos nexo entre la universidad y la empresa.

Nuestros productos tienen **presencia internacional**: Alemania, Argentina, Australia, Brasil, Colombia, EE. UU., España, Italia, México, Panamá, Paraguay, Perú, Portugal, Reino Unido, Rumanía,Venezuela.

Puedes desarrollar tu carrera profesional como analista, desarrollador o científico de datos en todos los sectores, siendo especialistas en:

Entorno Bancario | Entorno RR.HH. | Entorno Seguros | Entorno Digital | Entorno Energía | Entorno Salud | Entorno Inteligencia de Cliente

TECNOLOGÍA Y ANÁLISIS DE DATOS AL SERVICIO DE RR. HH.

HR ANALYTICS

El análisis de los datos de RR. HH. permite obtener información de valor para una mejor gestión del talento. El IIC aplica **analítica descriptiva y predictiva** para optimizar procesos de selección, predecir el absentismo o la rotación e identificar a los profesionales con más potencial, entre otros proyectos.

EVALUACIÓN DE COMPETENCIAS

Dentro de la **plataforma online eValue,** desarrollamos pruebas objetivas y fiables para evaluar las competencias transversales, el nivel de inglés o las motivaciones de candidatos y empleados. Además de tomar mejores decisiones, se obtienen **datos de calidad** para analizar, por ejemplo, sus necesidades de formación.

ANÁLISIS DE REDES ORGANIZACIONALES

Los proyectos AROS permiten analizar las relaciones de trabajo y las interacciones entre los profesionales. Representadas visualmente en un grafo, se pueden identificar **redes informales, referentes ocultos o cuellos de botella** en la organización, para emprender acciones de mejora.

www.linkedin.com/company/instituto-de-ingenier-a-del-conocimiento---iic

www.twitter.com/IIConocimiento

www.youtube.com/IIConocimiento

INNOVADATA

NUESTROS ASOCIADOS:

Santander | IBM | UAM Universidad Autónoma de Madrid

Instituto de Ingeniería del Conocimiento
C/ Francisco Tomás y Valiente, 11 EPS, Edificio B, 5ª planta UAM Cantoblanco.
28049 Madrid
http://www.iic.uam.es/empleo-iic/
rrhh@iic.uam.es
(+34) 91 497 2323

20
21

instituto
de ingeniería
del conocimiento

www.iic.uam.es

Instituto de Ingeniería del Conocimiento

C/ Francisco Tomás y Valiente, nº 11
Escuela Politécnica Superior (EPS),

Edificio B, 5ª planta
Universidad Autónoma de Madrid (UAM).

28049 Cantoblanco, Madrid

T. (+34) 91 497 23 23