**Gender Pay Gap**

**Katherine Pennacchio**

Para los resultados presentados a continuación se trabajó con un data set del gobierno del Reino Unido que muestra el reporte hecho por 1260 compañías sobre su brecha salarial de género.

Se debe considerar que la última data proporcionada es del 7 de enero de 2022 y que la fecha limite para la presentación de los reportes de cada compañía era el 31 de marzo y 5 de abril de 2022. Por lo tanto, los resultados no representan una visión de la realidad, solo una captura de la situación de las compañías que habían presentado hasta la fecha.

Otro problema que hay con el data set es que está construida con la información proporcionada por los mismos empleadores. El gobierno del Reino Unido no explica si se hace algún fact-checking de los datos y además muchos datos no se proporcionan. El 34,5% de las empresas no incluyó el URL de su sitio web y el 33% no proporcionó información sobre el % de diferencia media entre el pago de bonificación de hombres y mujeres.

Para calcular la brecha salarial de género de las empresas en el data set, se utilizó la columna llamada *DiffMedianHourlyPercent* que representa la diferencia porcentual media entre el salario por hora de hombres y mujeres. Es decir, la mediana. Se usó la mediana porque el valor medio del conjunto de datos refleja mejor la realidad que el promedio. El promedio puede verse afectado por valores muy bajos o muy altos.

En nuestro data set la brecha salarial favorece a los hombres por encima de las mujeres. Esto se ve en el elevado porcentaje de valores positivos, es decir, por encima de cero. 952 empresas de las 1260 que representa el 75,5%.

También se buscó cuáles eran las 10 compañías con la mayor brecha salarial a favor de los hombres. En primer lugar, se ubicó “Brand Energy & Infrastructure Services UK, LTD”, una empresa que proporciona servicios de andamios y equipos de acceso, revestimiento y pintura, limpieza, aislamiento y protección contra incendios a algunos de los clientes industriales más grandes del Reino Unido. Esta empresa cuenta con más de 1000 trabajadores.

En la tabla se puede ver como no existe una relación entre el número de trabajadores de una compañía y su brecha salarial. Otros cálculos realizados permiten concluir que la mediana de la brecha salarial de género es mayor en el número de empresas que no suministraron su cantidad de empleados.

Interfaz de usuario gráfica, Aplicación, Tabla, Excel

Descripción generada automáticamente

Las empresas que no proporcionaron su número de empleados eran en su escuelas, academias y hospitales.

Usando la mediana también se calculó la brecha salarial en todas las empresas. El resultado que se obtuvo es de 9.4.

Uno de los puntos a mejorar en el data set es que no tenemos la información sobre cuántos de los empleados son mujeres y cuántos son hombres. Por lo tanto, pone en duda la veracidad de los datos en el número de empresas que aseguran que su brecha salarial es cero. ¿Cómo sabemos cuántos de sus trabajadores son hombres y cuántos mujeres? Si todos sus trabajadores son mujeres, por ejemplo, la brecha salarial evidentemente sería cero.

En el data set son 131 empresas que reportan que su brecha salarial es cero. Solo 20 de ellas están en Londres. La mayoría está en Windsor England.

Siguiendo la línea de la ubicación, se calculó el promedio de la brecha salarial de género total de todas las empresas en Londres versus las que se encuentran fuera de Londres. El resultado es que la brecha salarial de género en Londres es mayor.

El data set proporciona también una columna de siccode que permite saber a qué categoría o qué tipo de empresa se está analizando. Sin embargo, hay un problema con esta columna. Puede haber varios siccodes en una celda por lo que los resultados en SQL no son confiables.

Por esta razón, se limpió la base y se calculó la mediana de la brecha salarial de género por siccode usando spreadsheet de Google drive porque en Excel no se consiguió hacer.

El resultado fue que el tipo de empresa con el promedio (mediana) de mayor brecha salarial es la que representa al siccode 85600. Es decir, servicios educativos de soporte. (Educational support services).

**Lista de queries usadas en la consulta del data set**

**1.¿Cuántas compañías están en el data set?**

SELECT DISTINCT employername

FROM public.gender\_pay\_gap

**RESULTADO: 1260**

**2-¿Cuántas de ellas enviaron la información luego del deadline?**

SELECT \*

FROM public.gender\_pay\_gap

WHERE datesubmitted > duedate

**RESULTADO: 0**

SELECT \*

FROM public.gender\_pay\_gap

WHERE submittedafterthedeadline = true

**RESULTADO: 0**

**3-¿Cuál fue la última fecha reportada de entrega de datos?**

SELECT MAX (datesubmitted)

FROM gender\_pay\_gap

**RESULTADO: 7 de enero de 2022**

**4- ¿Cuántas compañías no proporcionaron su URL?**

SELECT \*

FROM public.gender\_pay\_gap

WHERE companylinktogpginfo IS null

**RESULTADO 435**

**5.¿Qué medida de la brecha salarial contiene mucha data nula y no debe ser usada en el análisis?**

SELECT \*

FROM public.gender\_pay\_gap

WHERE diffmeanbonuspercent IS null

**RESULTADO 414**

**6. Consigue el promedio de la brecha salarial en todas las empresas**

SELECT percentile\_disc(0.5) within group (order by diffmedianhourlypercent)

FROM gender\_pay\_gap

**RESULTADO 9.4**

**7. ¿Cuáles son las 10 compañías con las mayores brechas salariales que favorecen a los hombres?**

SELECT DISTINCT currentname, MAX(diffmedianhourlypercent) AS largest\_pay\_gap,

FROM gender\_pay\_gap

GROUP BY currentname

ORDER BY 2 DESC

LIMIT 10

**8. Algunos filtros adicionales aplicados a esas compañías**

SELECT DISTINCT currentname, MAX(diffmedianhourlypercent) AS largest\_pay\_gap, employersize

FROM gender\_pay\_gap

GROUP BY currentname, employersize

HAVING employersize = '1000 to 4999'

ORDER BY 2 DESC

LIMIT 10;

**Otra opción**

SELECT DISTINCT currentname, MAX(diffmedianhourlypercent) AS largest\_pay\_gap, employersize

FROM gender\_pay\_gap

GROUP BY currentname, employersize

HAVING employersize = '5000 to 19,999'

ORDER BY 2 DESC

LIMIT 10;

**9. ¿Cuál es el promedio de la brecha salarial de género en Londres versus fuera de Londres?**

SELECT percentile\_disc(0.5) within group (order by diffmedianhourlypercent)

FROM gender\_pay\_gap

WHERE address ILIKE '%LONDON%';

**RESPUESTA 11.4**

SELECT percentile\_disc(0.5) within group (order by diffmedianhourlypercent)

FROM gender\_pay\_gap

WHERE address NOT ILIKE '%LONDON%';

**RESPUESTA 9.2**

**10. ¿Cuál es el promedio de la brecha salarial de género en Londres versus Birmingham?**

SELECT percentile\_disc(0.5) within group (order by diffmedianhourlypercent)

FROM gender\_pay\_gap

WHERE address ILIKE '%Birmingham%';

**RESULTADO 6.3**

**11. ¿Cuál es el promedio de la brecha salarial de género en escuelas?**

SELECT percentile\_disc(0.5) within group (order by diffmedianhourlypercent)

FROM gender\_pay\_gap

WHERE currentname ILIKE '%school%';

**RESULTADO 23.7**

**12. ¿Cuál es el promedio de la brecha salarial de género en bancos?**

SELECT percentile\_disc(0.5) within group (order by diffmedianhourlypercent)

FROM gender\_pay\_gap

WHERE currentname ILIKE '%bank%';

**RESULTADO 32.1**

**13. ¿Alguna relación entre el número de empleados de una compañía y su promedio de brecha salarial?**

|  |  |
| --- | --- |
|  |  |

SELECT employersize, percentile\_disc(0.5) within group (order by diffmedianhourlypercent)

FROM gender\_pay\_gap

GROUP BY employersize

ORDER BY 2 DESC

RESULTADO:

La media del gender pay gap mayor está en las empresas que no suministraron el número de empleados. Luego está los de 20.000 o más que se ubican en 10.6 a favor de los hombres. No se podría decir que hay una relación entre el número de empleados de la compañía y el gap.

**PREGUNTAS ADICIONALES**

**14. ¿Qué empresas no proporcionaron los datos del número de trabajadores?**

SELECT \*

FROM gender\_pay\_gap

WHERE employersize ILIKE 'Not Provided'

**RESULTADO: 35  
Son en su mayoría escuelas, academias, hospitales.**

**15. ¿En cuántas empresas el pay gap favorece a las mujeres?**

SELECT \*

FROM gender\_pay\_gap

WHERE diffmedianho urlypercent < 0

**RESULTADO 177 de 1260 (total)**

**16. ¿En cuántas empresas el pay gap favorece a las hombres?**

SELECT \*

FROM gender\_pay\_gap

WHERE diffmedianhourlypercent > 0

**RESULTADO 952 de 1260 (total)**

**17. ¿En cuántas empresas no hay un gap?**

SELECT \*

FROM gender\_pay\_gap

WHERE diffmedianhourlypercent = 0

**RESULTADO: 131 de 1260 (total)**

**18. ¿Cuántas de estas de la pregunta anterior están en Londres?**

SELECT \*

FROM gender\_pay\_gap

WHERE diffmedianhourlypercent =0 AND address ilike '%london%'

20 de ellas están en Londres. Es decir, la mayoría está fuera de Londres. La mayoría está en Windor England.

SELECT postcode, count (postcode)

FROM gender\_pay\_gap

WHERE diffmedianhourlypercent =0

GROUP BY postcode

ORDER BY 2 DESC

**19.¿Cuál es el siccode (tipo de empresa) predominante en las que no tienen brecha salarial?**

SELECT DISTINCT siccodes, COUNT (siccodes)

FROM gender\_pay\_gap

WHERE diffmedianhourlypercent = 0

GROUP BY siccodes

ORDER BY 2 desc

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

**RESULTADO: 81210** pertenece a General cleaning of buildings. Pero, hay un problema con esta columna de siccode. Están varias siccodes juntos en una celda.

Por esta razón, se limpió la base y se calculó la mediana de la brecha salarial de género por siccode usando spreadsheet de Google drive porque en Excel no se consiguió hacer.

El resultado fue que el tipo de empresa con el promedio (mediana) de mayor brecha salarial es la que representa al siccode 85600. Es decir, servicios educativos de soporte. (Educational support services)