



INSTITUTO DE FORMACIÓN TÉCNICA SUPERIOR N°29

**TECNICATURA SUPERIOR EN DESARROLLO
DE SOFTWARE**

**ESTADÍSTICA Y PROBABILIDADES PARA EL DESARROLLO DEL
SOFTWARE**

ANÁLISIS DE LA ENCUESTA DE SUELDOS IT - SYSARMY 2025

COMISIÓN: 'D'

PROFESORA: DANIELA PARADA

APELLIDO Y NOMBRE: FACAL, XIMENA BELÉN
LUGONES, MELINA
PINTO, RODRIGO
SKAARUP, MARA LUZ

ÍNDICE

INTRODUCCIÓN (3)

DESCRIPCIÓN DE LOS DATOS (4)

VARIABLES QUE SE USARON (5)

DATOS DE LA ENCUESTA (7)

LIMPIEZA DE DATOS (7)

EXPLORACIÓN INICIAL DE DATOS (8)

HIPÓTESIS Y ANÁLISIS (15)

CONCLUSIONES (21)

1.INTRODUCCIÓN

En este informe se analiza la encuesta de sueldos publicada por Sysarmy correspondiente al primer trimestre de 2025. El objetivo es explorar la relación entre el salario mensual y variables como el modo de trabajo o la moneda de cobro. Se replican, amplían y actualizan análisis realizados previamente con los datos del año 2020.

La Encuesta Sysarmy es un estudio anual consolidado que, desde 2014, se ha erigido como un referente fundamental para comprender la dinámica salarial y las condiciones laborales dentro de la comunidad tecnológica argentina.

La encuesta se enfoca en la comunidad tecnológica argentina, abarcando un amplio espectro de roles dentro del ámbito de TI. Su objetivo es la recopilación de datos sobre salarios y beneficios. Esto incluye no solo el sueldo base, sino también información detallada sobre bonificaciones, proyecciones de aumentos, horas de trabajo, y una variedad de beneficios adicionales que son relevantes para la compensación total de los profesionales.

Toda la información recopilada se hace pública a través del blog oficial de Sysarmy, garantizando que cualquier persona pueda acceder y analizar los datos de forma libre. Proporciona una visión general clara de las tendencias salariales y las condiciones laborales predominantes, lo que es crucial tanto para los profesionales como para las empresas del sector. Finalmente, cabe destacar que la Encuesta funciona como una plataforma colaborativa. Los participantes contribuyen activamente con sus propios datos, y a cambio, se benefician de la rica información compartida por el conjunto de la comunidad, fomentando un ciclo virtuoso de transparencia y crecimiento colectivo.

En este informe se presenta un análisis de los resultados de la encuesta de sueldos realizada por la comunidad Sysarmy en el primer trimestre de 2025. Esta encuesta es una fuente confiable y colaborativa que releva información salarial en el sector tecnológico. El objetivo es analizar cómo distintos factores impactan en los salarios mensuales brutos.

2.DESCRIPCIÓN DE LOS DATOS

Origen del Dataset

La base de datos utilizada corresponde a la Encuesta de Sueldos del primer trimestre de 2025, publicada por Openqube, un proyecto impulsado por la comunidad Sysarmy, que releva desde 2014 información sobre condiciones laborales en el sector IT.

La encuesta es colaborativa y anónima, respondida voluntariamente por trabajadores del ámbito tecnológico (desarrolladores, analistas, testers, sysadmins, data scientists, etc.) en su mayoría de Argentina, pero también de otros países de habla hispana.

3.VARIABLES QUE SE USARON

El dataframe que se creo contiene filas y columnas, las filas corresponde a cada individuo que contestó la encuesta y las columnas son variables; las cuales detallamos en el siguiente cuadro:

Tipo de Variable	Variable
Cualitativas nominales	<ul style="list-style-type: none">• donde_estas_trabajando• tipo_de_contrato• pagos_en_dólares• a_que_esta_atado_el_bono• modalidad_de_trabajo• plataformas_que_utilizas_en_tu_puesto_actual• lenguajes_de_programacion_o_tecnologias_que_utilices_en_tu_puesto_actual• frameworksherramientas_y_librerias_que_utilices_en_tu_puesto_actual• bases_de_datos• qa_testing• opcionalen_que_empresaorganizacion_estas_trabajando• carrera• institución_educativa• genero• sueldo_dolarizado• trabajo_de

Tipo de Variable	Variable
Cualitativas ordinales	<ul style="list-style-type: none"> • dedicacion • recibis_algun_tipo_de_bono • tuviste_actualizaciones_de_tus_ingresos_laborales_durante_2024 • como_consideras_que_estan_tus_ingresos_laborales_comparados_con_el_semestre_anterior • que_tan_conforme_estas_con_tus_ingresos_laborales • estas_buscando_trabajo • cantidad_de_personas_en_tu_organizacion • la_recomendas_como_un_buen_lugar_para_trabajar • que_tanto_estas_usando_copilotchatgpt_u_otras_herramientas_de_ia_para_tu_trabajo • maximo_nivel_de_estudios • estado • seniority
Cuantitativas discretas	<ul style="list-style-type: none"> • años_de_experiencia • antigüedad_en_la_empresa_actual • años_en_el_puesto_actual • cuantas_personas_tenes_a_cargo • si_trabajas_bajo_un_esquema_hibrido_cuantos_dias_a_la_semana_vas_a_la_oficina • tengo_edad
Cuantitativas Continuas	<ul style="list-style-type: none"> • ultimo_salario_mensual_o_retiro_bruto_en_pesos_argentinos • ultimo_salario_mensual_o_retiro_netto_en_pesos_argentinos • si_tu_sueldo_esta_dolarizado_cual_fue_el_ultimo_valor_del_dolar_que_tomaron • de_que_fue_el_ajuste_total_acumulado

4.LIMPIEZA DE DATOS

- Se excluyeron edades mayores a 70 años.
- Se filtraron casos con información nula en variables clave.
- Herramientas: Python (pandas, seaborn, matplotlib, scipy).
- Criterios de limpieza: se excluyeron outliers extremos (> P98), errores obvios en edad/salario.

5.DATOS DE LA ENCUESTA:

Participaron en la encuesta un total de 5.196 personas, respondiendo alrededor de 50 preguntas, de las cuales podemos mostrar los datos más relevantes. Nuestro Dataset tiene 5.196 filas y 51 columnas.

El archivo contiene miles de registros (respuestas individuales a la encuesta).

Cada fila representa a una persona que trabaja en IT y completó el formulario.

El archivo está en formato CSV con delimitador por coma y los valores están parcialmente normalizados (algunas columnas como technologies o salary_currency pueden tener formatos libres o mixtos).

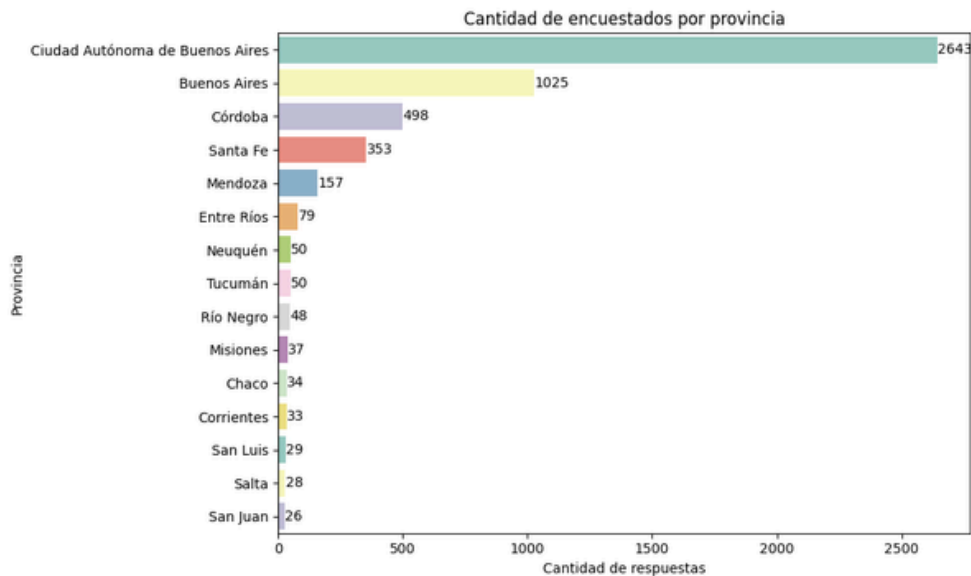
Algunas columnas presentan datos faltantes (NaN), por lo que fue necesario tratarlos para evitar sesgos.

6. EXPLORACIÓN INICIAL DE DATOS

1. Distribución geográfica por provincia

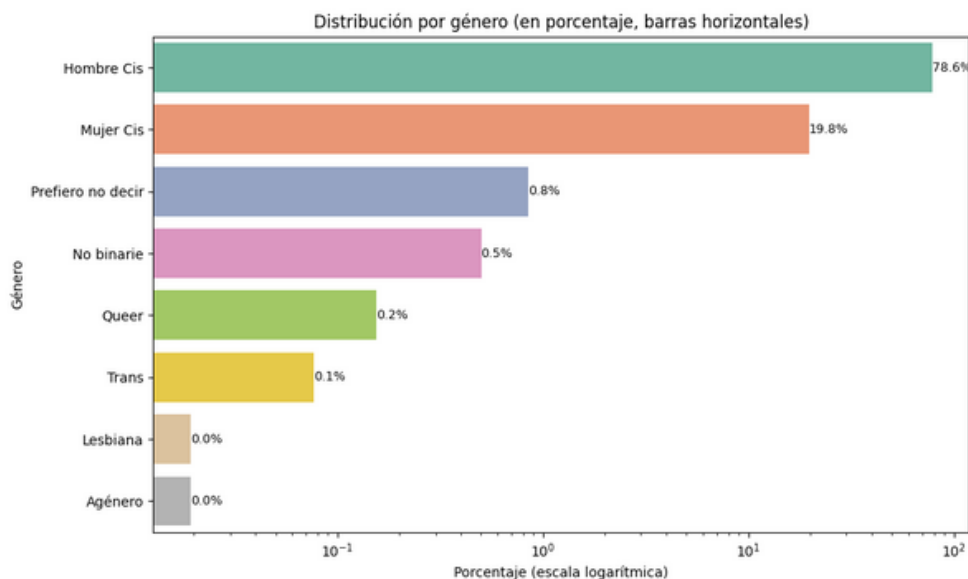
El gráfico muestra que la mayoría de las respuestas provienen de la Ciudad Autónoma de

Buenos Aires (2643), seguida por la Provincia de Buenos Aires, Córdoba y Santa Fe. Esto refleja el alto nivel de centralización del sector IT en el AMBA.



2. Distribución porcentual por género

La distribución de género es altamente desequilibrada. Casi el 79% se identifican como "Hombre", seguido por "Mujer" con un 19.8%. El resto de las identidades tienen muy baja representación, menor al 1%, lo que implica que los análisis estadísticos por género podrían estar sesgados por el desbalance muestral, particularmente en comparaciones salariales.

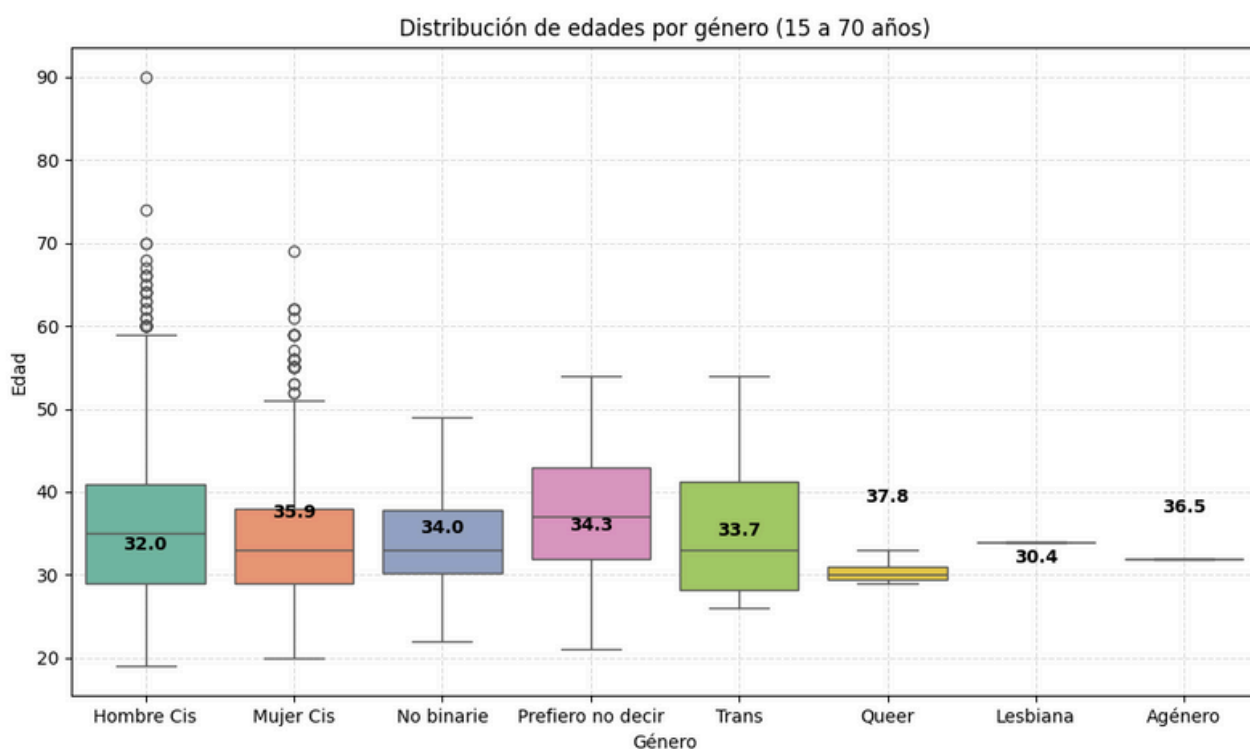


3. Distribución por edad y género:

este grafico nos muestra la distribución por edad y género, el mayor porcentaje son hombres, indicando que aún hoy hay pocas mujeres en IT, y mucho más en los otros generos. La mayoría de los encuestados tienen entre 25 y 40 años, siendo el promedio general cercano a los 33 años. El grupo "Hombre" es el más numeroso y tiene una mediana de edad de 32 años, mientras que "Mujer" tiene una mediana algo superior (35.9 años). También se observan algunos outliers en los extremos de cada grupo. Estos datos reflejan una población joven-adulta en edad laboral activa, con predominancia de perfiles masculinos.

Edad promedio: 35.6

Rango de edad: 19 a 70



Medidas de interés

n = 5196

Media = 2728422.7280

σ (poblacional) = 1780070.0646

s (muestral) = 1780241.3816

Σx = 14176884494.67

Σx^2 = 5.514483633038848e+16

Mín = 150000.0

Q1 = 1500000.0

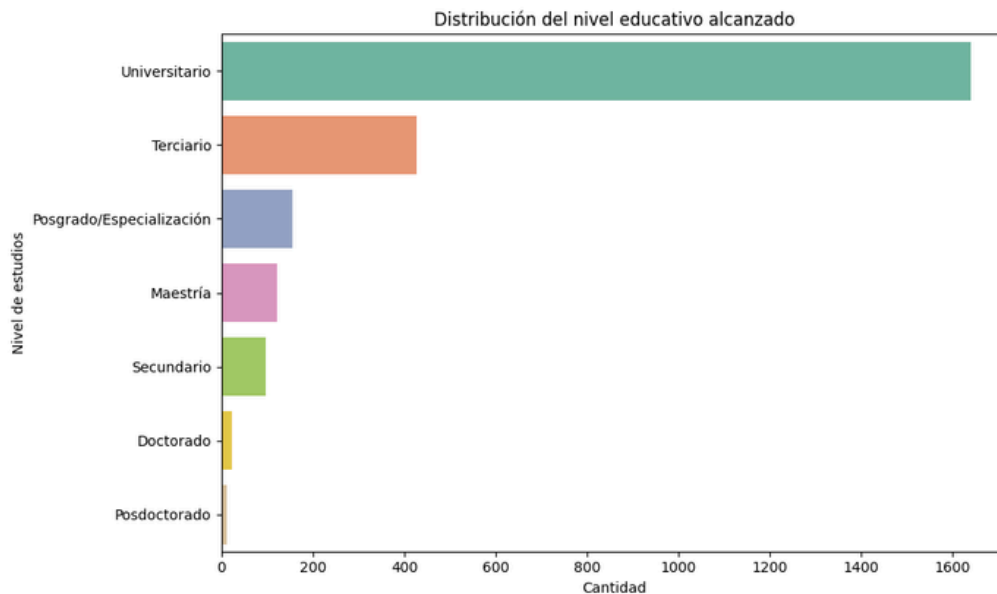
Mediana = 2290000.0

Q3 = 3500000.0

Máx = 13500000.0

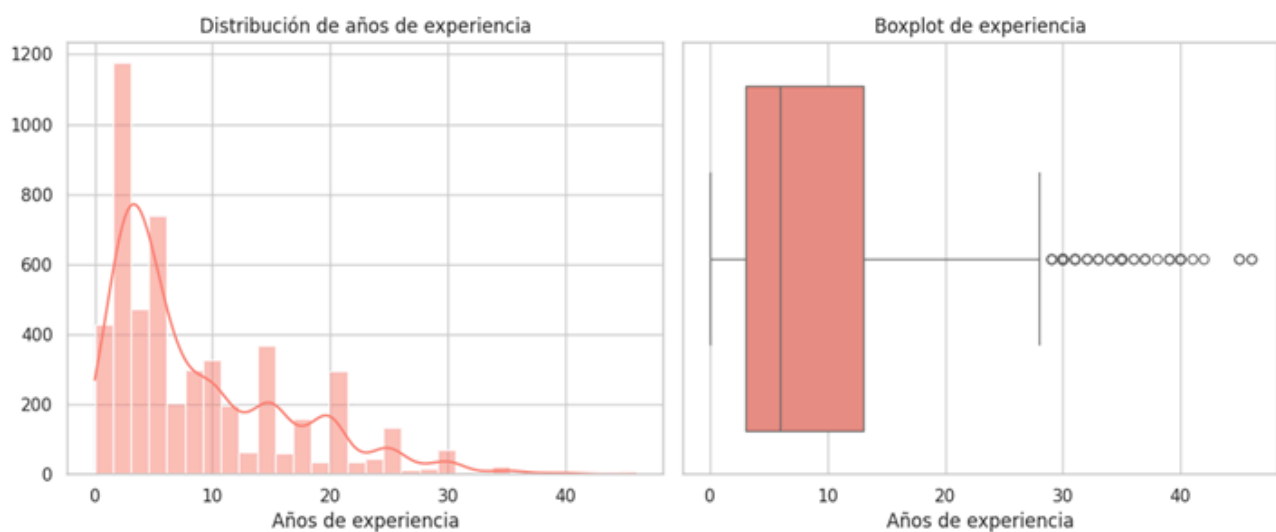
4. Nivel educativo alcanzado

Se observa una fuerte mayoría de encuestados con estudios universitarios completos o en curso, seguidos por quienes cursaron nivel terciario. Esto coincide con el perfil técnico que demanda el sector. También existen niveles más altos (posgrado, doctorado), pero en menor proporción. Este dato será relevante si se quiere controlar por formación al analizar brechas salariales.



5. Distribución de años de experiencia

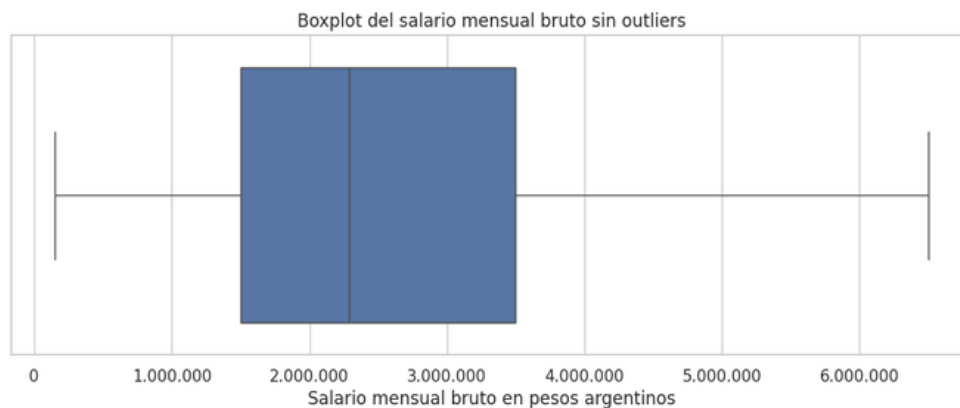
La mayoría de las personas encuestadas tiene entre 2 y 10 años de experiencia laboral. El histograma revela una fuerte concentración en los primeros años de carrera, mientras que el boxplot muestra una dispersión significativa, con valores atípicos hacia los extremos (casos de 40+ años).



6. Promedio del salario mensual

Este gráfico muestra la distribución del salario bruto mensual luego de haber removido valores extremos (outliers). La media se encuentra cerca de los 2.5 millones de pesos, y los cuartiles sugieren una distribución razonablemente simétrica sin valores atípicos que distorsionen el análisis central.

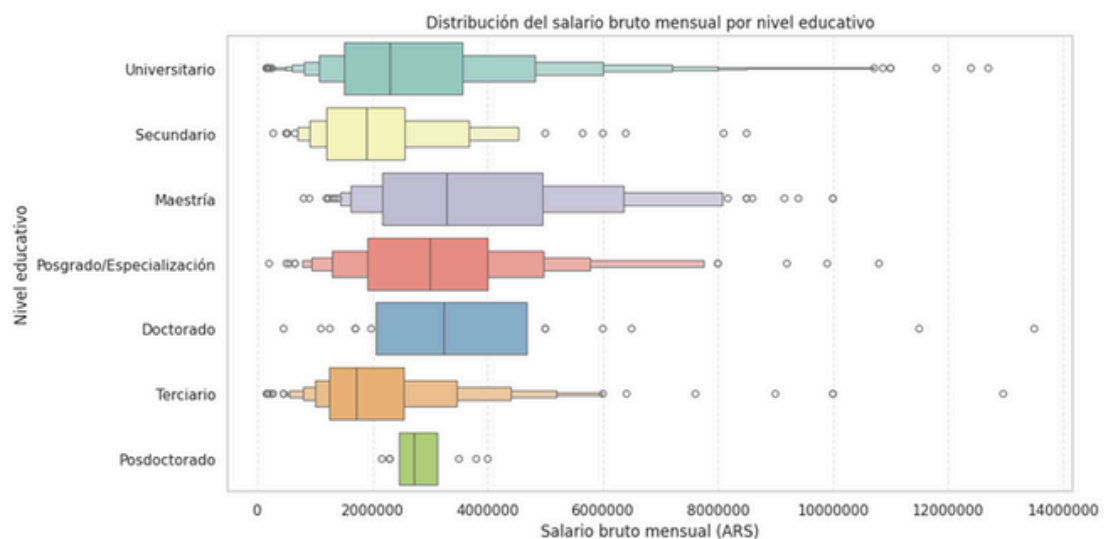
Esto ayuda a establecer un rango salarial típico sobre el cual construir comparaciones válidas para otras variables.



7. Distribución del salario por nivel educativo

Se visualiza cómo varía el salario según el nivel educativo alcanzado. Se observa que niveles más altos como Doctorado, Maestría y Posgrado tienden a tener mayores medianas salariales. No obstante, hay mucha superposición entre los grupos y una amplia dispersión.

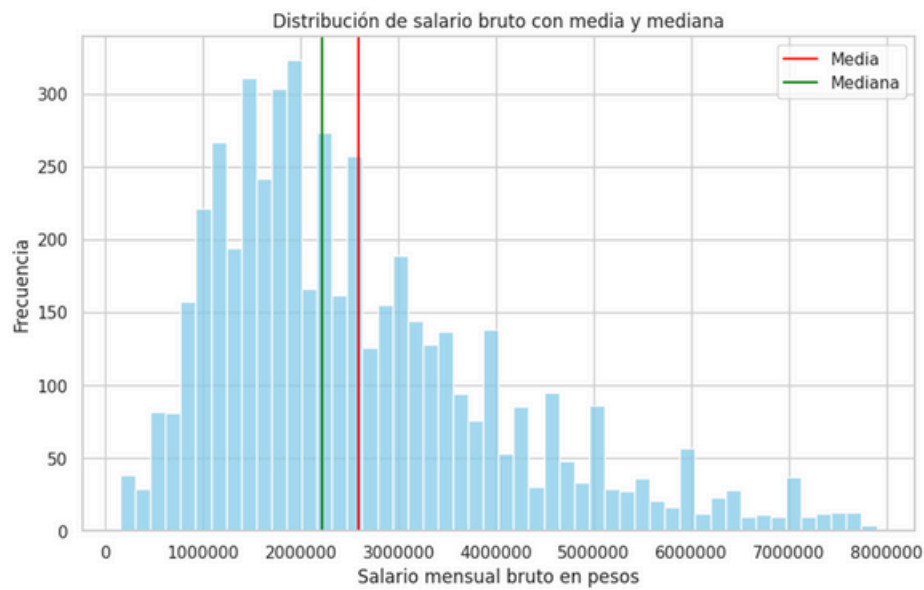
Esto puede indicar que el nivel educativo tiene cierto impacto en el salario, pero también que no es el único factor determinante, por ejemplo, el rol podría pesar más.



8. Distribución del salario bruto mensual

Presenta una clara asimetría hacia la derecha, como es habitual en salarios. La mayoría de los encuestados perciben entre 1.500.000 y 3.000.000 mensuales, con una media superior a la mediana, lo que indica que los valores altos están empujando el promedio hacia arriba.

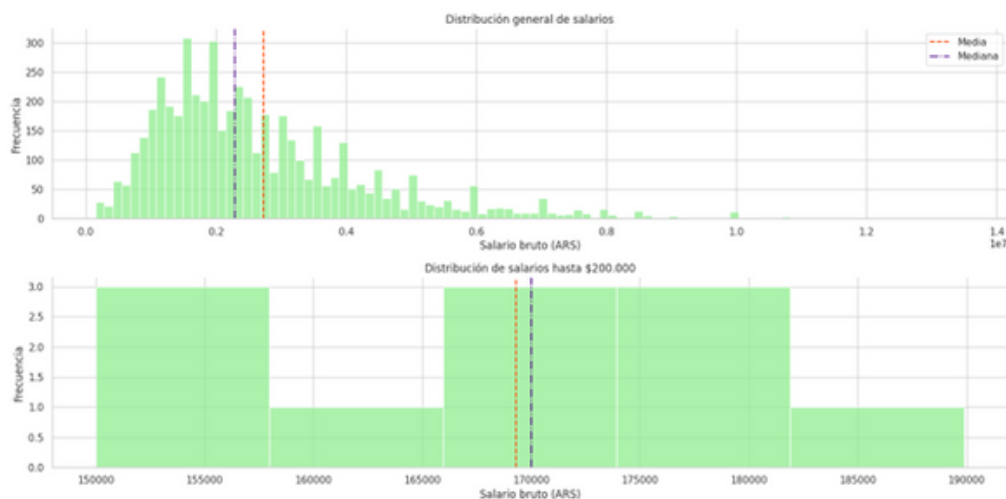
Los histogramas muestran que existe una cola larga de salarios muy altos (outliers), por lo que se consideró filtrar o analizar el percentil 98 para evitar sesgos en la media. Esta observación sustenta la necesidad de explorar diferencias según modalidad de trabajo y tipo de moneda, ya que podrían explicar la dispersión.



9. Histograma del salario bruto con y sin filtro

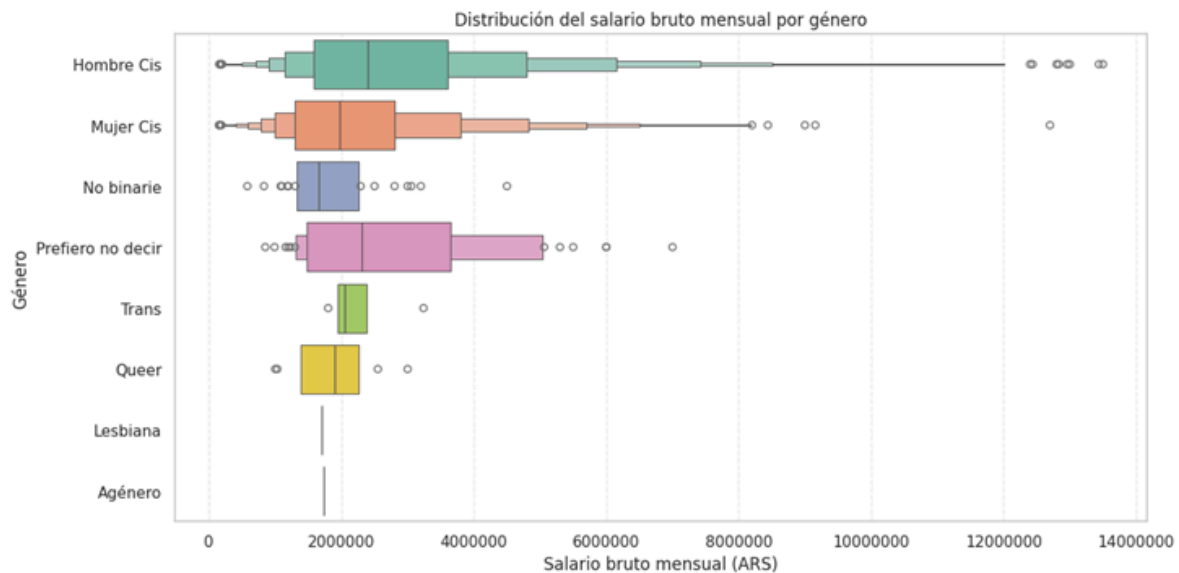
Este gráfico muestra la distribución del salario bruto con y sin un recorte superior (outliers). En ambos casos, la media es mayor a la mediana, lo que sugiere una distribución sesgada a la derecha (asimetría positiva). Esto es típico en datos salariales, donde hay pocos valores muy altos que elevan el promedio.

Este patrón valida el uso de la mediana como una métrica más representativa que la media para análisis de comparación.



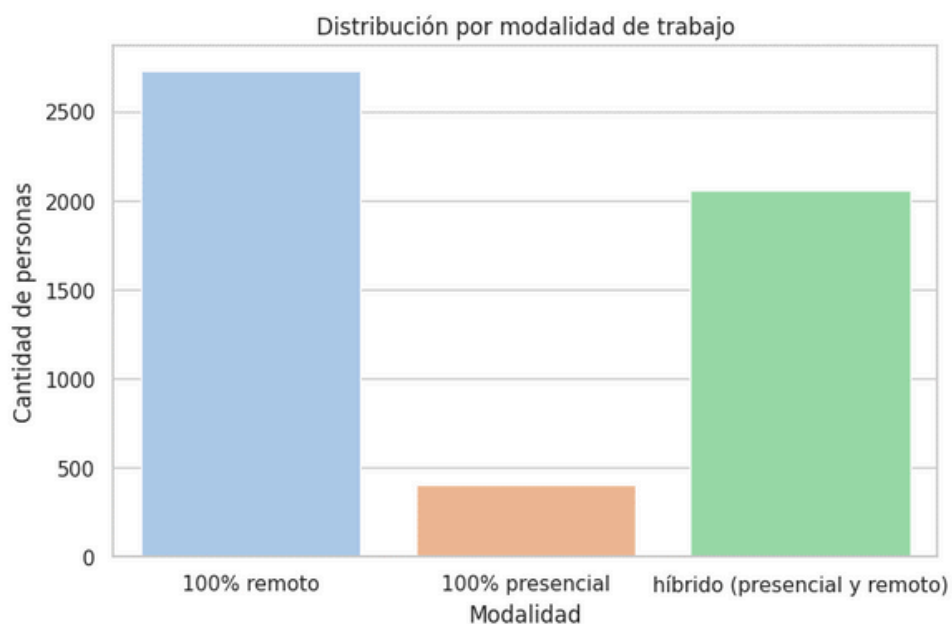
10. Distribución del salario bruto mensual por género

Vemos la comparación de la distribución del salario entre diferentes identidades de género. Se aprecia que Hombres y Mujeres concentran la mayor cantidad de datos, pero los salarios de los Hombres tienen una mediana y una dispersión superiores. Otros grupos como personas no binaries, queer o trans muestran una distribución con menor mediana y más dispersión, aunque con menos datos disponibles.



11. Personas por modalidad de trabajo (presencial, remoto, híbrido)

Este gráfico nos muestra la cantidad de individuos en función de la modalidad de trabajo. Los resultados indican que son más las personas que trabajan 100% remoto. Seguidos por los trabajadores de modo híbrido y por último los empleados 100% presenciales quienes son los menos.



12. Salario bruto por modalidad de trabajo

Gráfico Presencial:

Tiene el salario bruto más bajo en promedio y mediana.

La mayoría de los valores se concentran por debajo de los 2 millones.

También hay menos dispersión y menos outliers con respecto a las otras modalidades.

Gráfico Híbrido:

Presenta un salario más elevado que el presencial, tanto en media como en mediana.

Mayor dispersión de datos, con sueldos que llegan a los 12 millones.

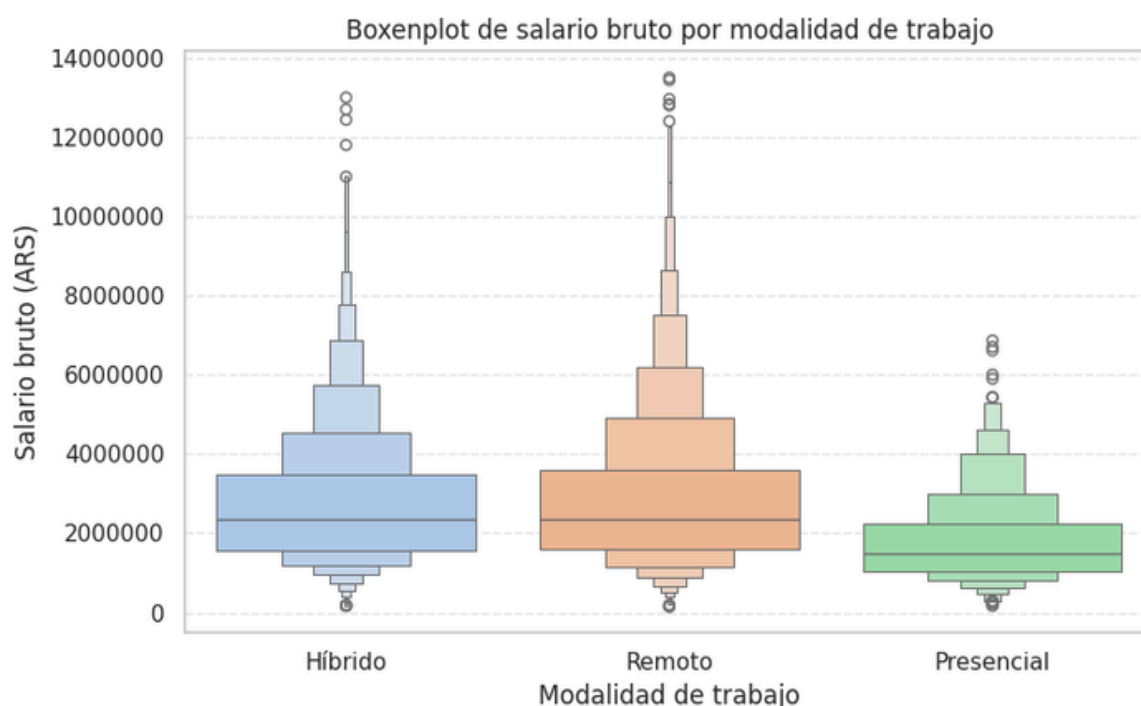
Hay más variabilidad y una distribución más extendida, lo cual indica heterogeneidad en los puestos o roles.

Gráfico Remoto:

Tiene la mediana y la dispersión más altas de las tres modalidades.

La base de la caja y sus "escalones" están más arriba, lo que indica que se gana más trabajando remoto.

También presenta más outliers altos, con algunos salarios que superan los 13 millones.



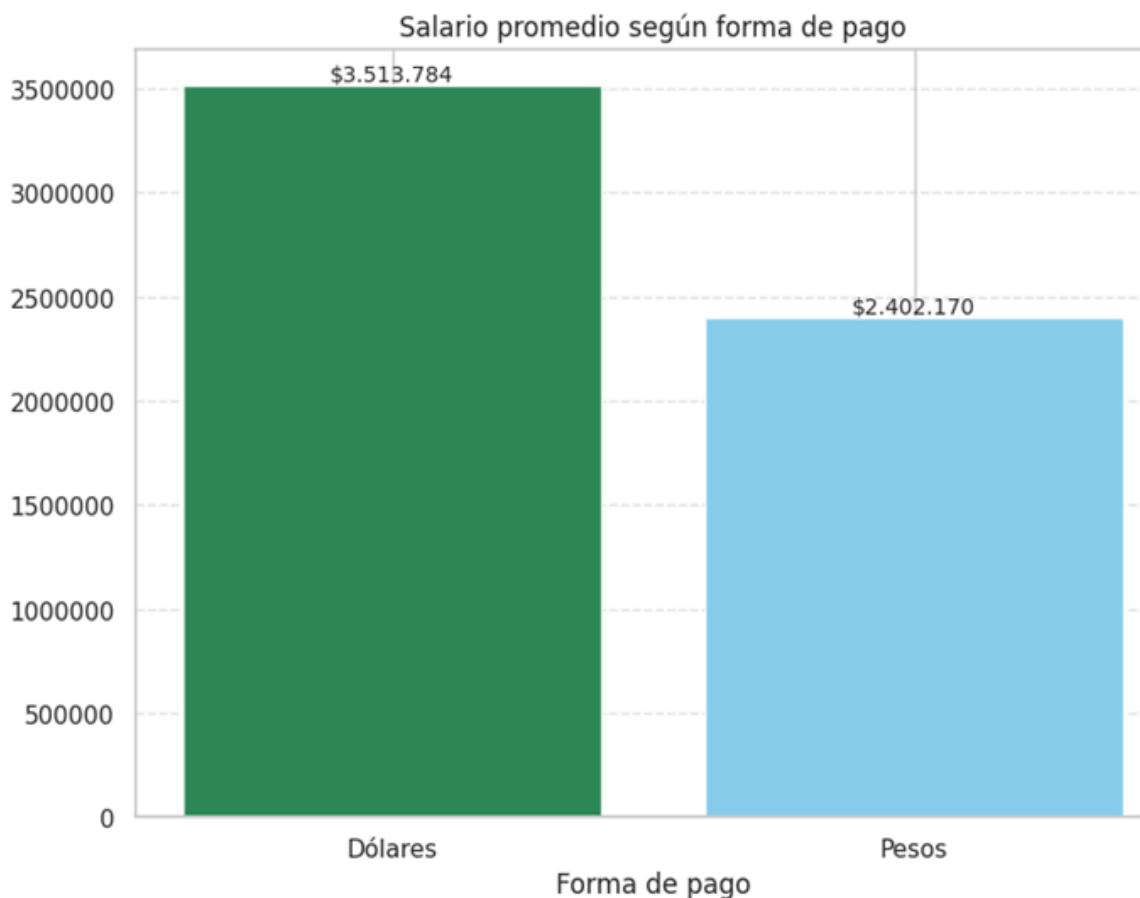
7. HIPÓTESIS Y ANÁLISIS

A partir de esta exploración, se identificaron tres ejes clave que queremos investigar:

- 1) Tipo de moneda: ¿la fuerte dispersión salarial podría estar influenciada por la dolarización de parte de los sueldos?
- 2) Modalidad de trabajo: ¿se presume que quienes trabajan de forma remota o híbrida tienen mayores ingresos?
- 3) Área tecnológica: ¿el rol profesional puede ser determinante en la remuneración?

Salario promedio según forma de pago (pesos vs dólares)

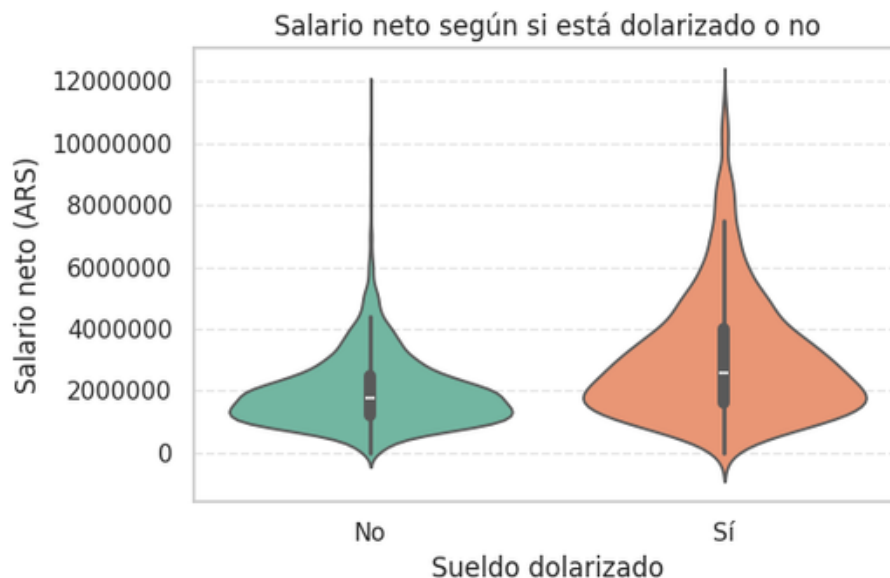
Este gráfico compara el salario bruto promedio entre quienes cobran en pesos argentinos y quienes lo hacen en dólares. Se observa que los salarios dolarizados tienen una diferencia significativa, siendo aproximadamente 46% más altos que los salarios en pesos. Esto sugiere que la moneda en la que se percibe el ingreso tiene un impacto fuerte sobre el nivel salarial expresado en pesos. Podría estar relacionado con puestos en empresas extranjeras o freelancing.



Salarios netos promedios según si están dolarizados o no

En este gráfico podemos observar que las personas con sueldos dolarizados ganan, en promedio, más que las que no..

Existe más variabilidad y potencialmente más desigualdad en los sueldos dolarizados.

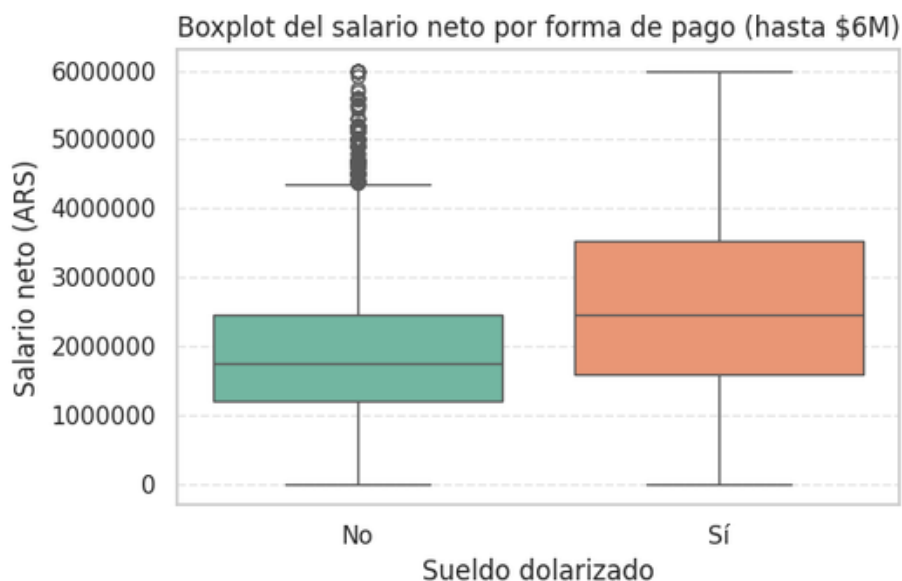


Salarios netos promedios por forma de pago

Las personas que cobran en dólares tienen un salario neto promedio más alto y también una mayor variabilidad.

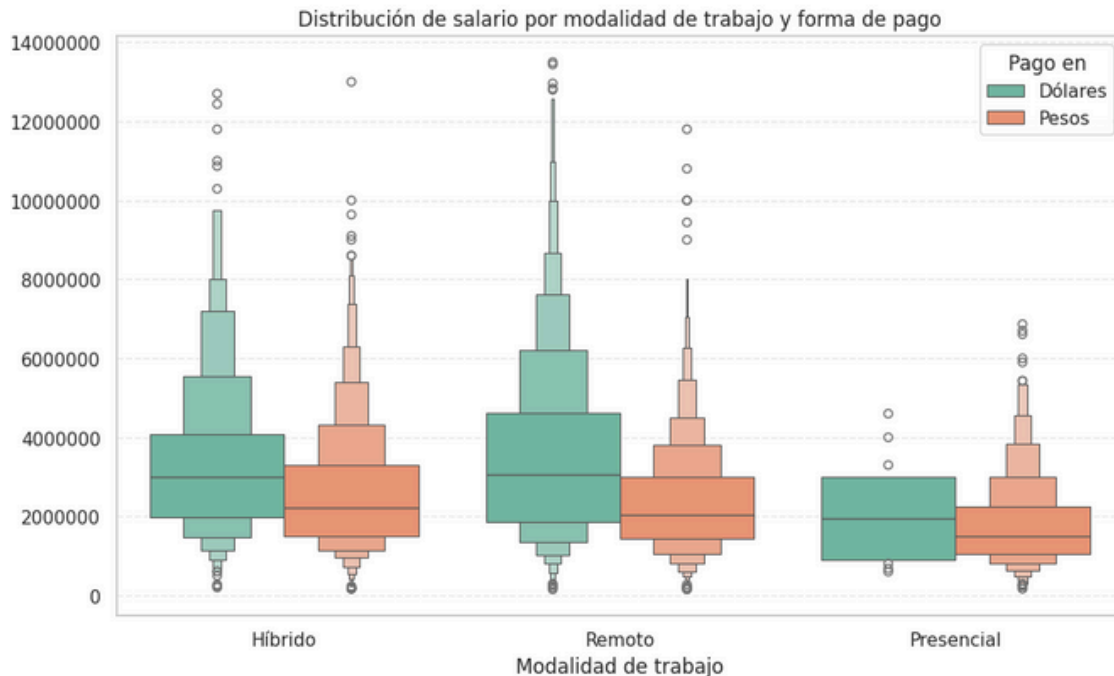
La mediana y el rango intercuartílico más altos en el grupo Dolarizado respaldan la hipótesis de que la dolarización influye positivamente en los ingresos.

El grupo No dolarizado tiene una distribución más concentrada y sueldos más bajos en promedio y general.



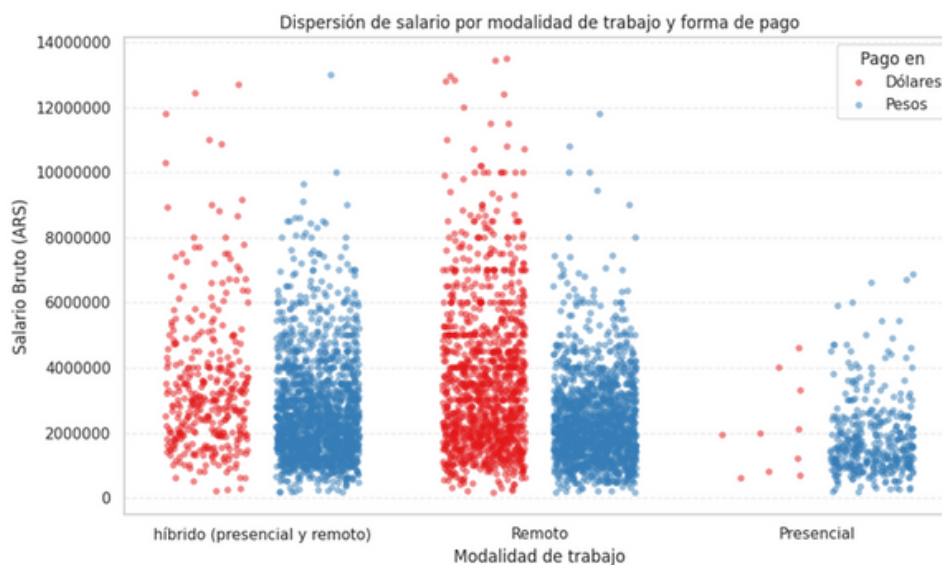
Distribución de salario por modalidad de trabajo y forma de pago

Los pagos en dólares (color verde) tienen mayor mediana, mayor rango intercuartílico y más outliers altos, especialmente en modalidad remota. Los pagos en pesos (color salmón) tienen menor variabilidad y menor nivel general de salarios.



Dispersión de salario por modalidad de trabajo y forma de pago

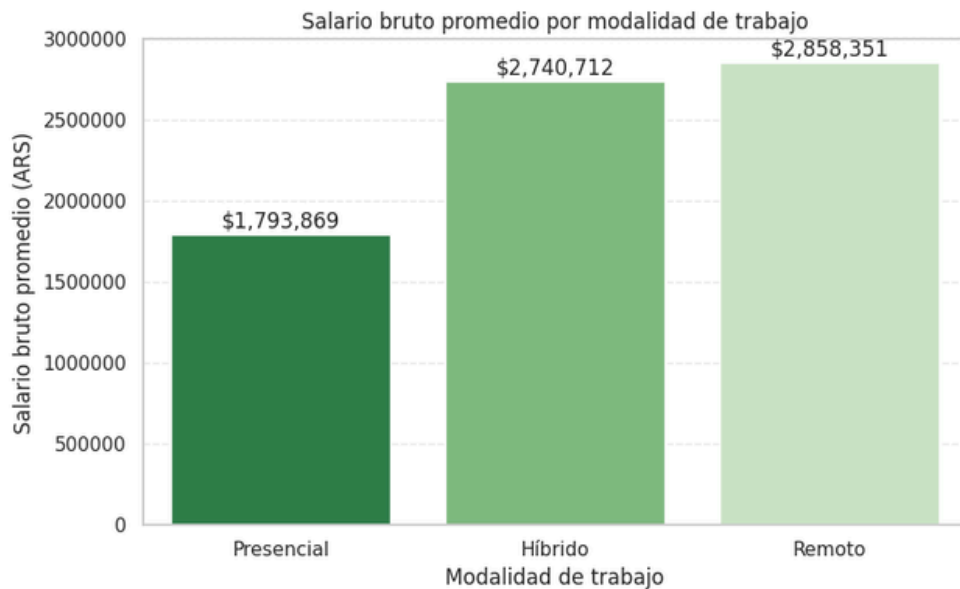
Los salarios en dólares (puntos rojos) son consistentemente más altos que los salarios en pesos (puntos azules), con mayor dispersión y presencia de valores extremos. Además, la modalidad remota aparece asociada a mayores niveles salariales, especialmente entre quienes cobran en dólares. La modalidad 100% presencial concentra los salarios más bajos.



Salario bruto promedio por modalidad de trabajo

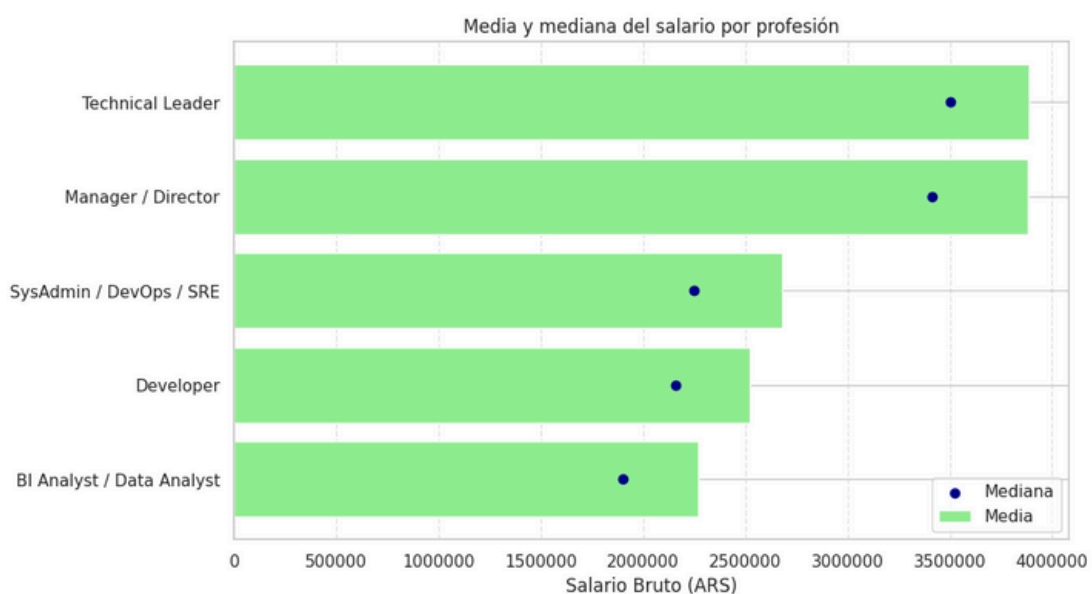
Este gráfico de barras muestra el salario bruto promedio según la modalidad de trabajo declarada por los encuestados.

Las personas que trabajan en modalidad remota tienen el promedio salarial más alto \$2.85 millones. Le sigue la modalidad híbrida con \$2.74 millones. La modalidad presencial se ubica muy por debajo, con un salario promedio cercano a los \$1.79 millones.



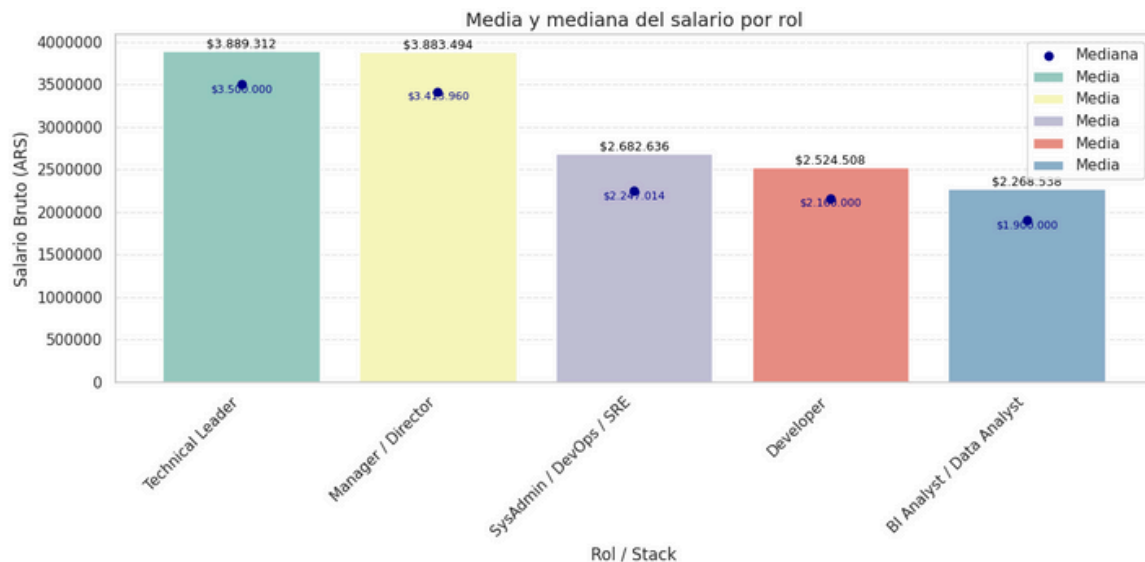
Media y mediana del salario por profesión (horizontal)

Technical Leaders y Managers ganan más que Developers o Analistas. Esta visualización resalta la diferencia entre media y mediana, importante para detectar asimetrías y justificar el análisis con ambas métricas.



Media y mediana del salario por rol

En este gráfico se comparan los salarios medios y medianos por distintos roles del sector IT. Se evidencia que los puestos de mayor jerarquía como Manager/Director y Technical Leader tienen los sueldos más altos, superando los 3,8 millones de pesos mensuales. A medida que se descende en la jerarquía técnica, los sueldos también disminuyen: desarrolladores y analistas de datos tienen salarios más bajos y con menor mediana.



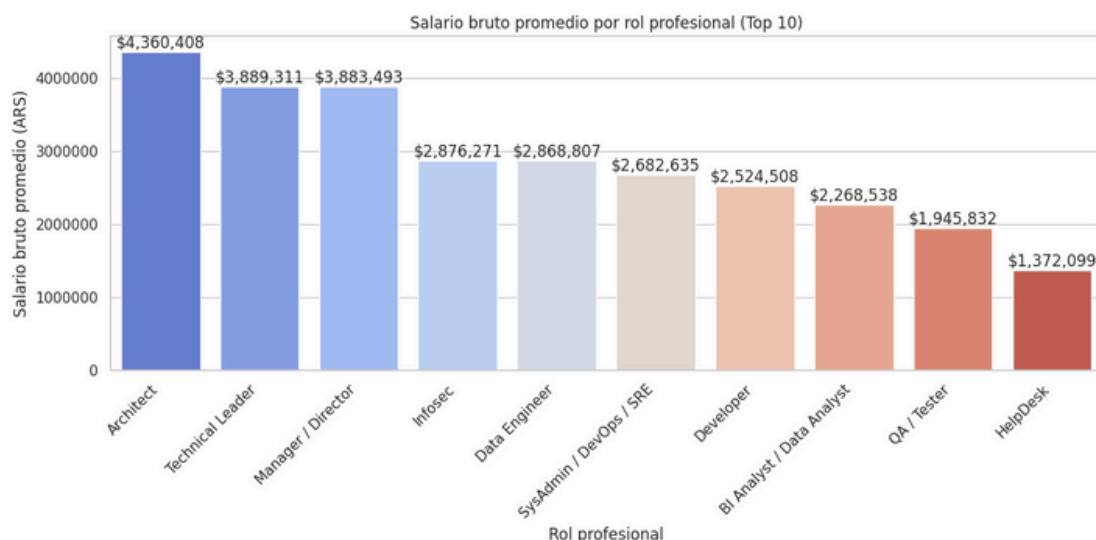
Salario bruto promedio por rol profesional

Este gráfico de barras horizontales muestra el salario bruto promedio de los 10 roles profesionales mejor remunerados según la encuesta. Los valores están expresados en pesos argentinos (ARS).

Architect, Technical Leader y Manager / Director son los tres roles mejor pagos, con salarios promedio superiores a \$3.8 millones.

HelpDesk y QA / Tester tienen los ingresos más bajos del Top 10, con HelpDesk por debajo de \$1.4 millones.

Los roles relacionados con infraestructura (DevOps/SRE), análisis de datos y desarrollo también se posicionan bien, pero con mayor variabilidad.



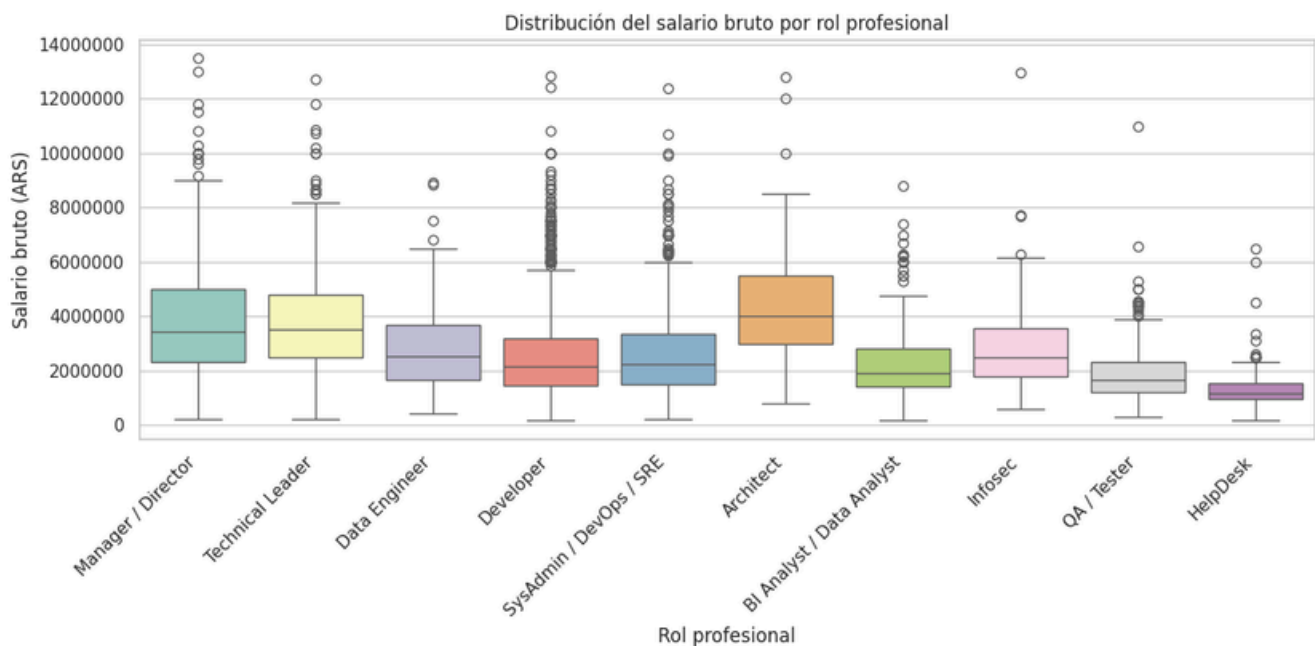
Distribución del salario bruto por rol profesional

Este gráfico de boxplot complementa al anterior mostrando la dispersión salarial para cada rol profesional.

Roles como Manager / Director, Architect y Technical Leader tienen tanto medianas altas como mayor número de outliers (valores extremos).

Developer, QA/Tester, BI Analyst/Data Analyst y HelpDesk muestran una mediana más baja y menor rango intercuartílico, lo que indica menor dispersión salarial.

La amplitud en roles técnicos como SysAdmin/DevOps/SRE o Data Engineer sugiere diversidad de condiciones según experiencia o empresa.



Desviación estándar Sueldo_dolarizado

False 1,364,182.47

True 2,100,513.26

- False = No dolarizado:
- La desviación estándar es 1.364.182,47
- Significa que los sueldos de este grupo varían en promedio $\pm 1,36$ millones respecto a su media. Es una dispersión moderada.
- True = Dolarizado:
- La desviación estándar es 2.100.513,26
- Este grupo tiene una variabilidad mayor. Los sueldos dolarizados fluctúan mucho más: algunos ganan bastante más que otros.

8. CONCLUSIONES

- **Reflexión final**

El análisis exploratorio nos permitió comprender como se distribuye el salario promedio, cuáles son las profesiones de IT mejor pagas. Como también que aún hoy los puestos de IT están ocupados mayormente por hombres. Se pudo ver las medias y medianas de los salarios, y otras variables como la cantidad por género, edades promedios, etc.

Pudimos trabajar con los distintos tipos de variables, convertir datos y filtrar. Todo este proceso nos ayudó a formular nuestras hipótesis de manera sólida para futuros informes.

A partir del análisis de la muestra de 5196 casos, observamos que el salario bruto mensual promedio es de aproximadamente \$2.728.423, mientras que la mediana es de \$2.290.000. Esta diferencia entre ambos valores me sugiere que hay cierta asimetría en la distribución, probablemente por la presencia de sueldos muy altos que elevan el promedio. Al mirar el histograma, notamos que la mayoría de los salarios se concentran entre \$1.500.000 y \$3.500.000, lo cual coincide con los valores del primer y tercer cuartil (Q1 y Q3).

En cuanto a la variabilidad, calculamos una desviación estándar muestral de alrededor de \$1.780.241. Esto indica que, en promedio, los sueldos se alejan bastante de la media. Al observar el diagrama de caja y bigotes, nos damos cuenta de que hay varios valores atípicos que se distancian mucho del resto, lo que probablemente hace que la desviación estándar sea más alta. Si no tuviera esos valores extremos, creo que la dispersión sería más baja.

También noté que la caja del diagrama está más cargada hacia la izquierda, lo que me lleva a pensar que los sueldos más altos están más dispersos que los más bajos. El 50% central de los datos (entre Q1 y Q3) se encuentra en un rango de 2 millones de pesos, lo que me dio una idea clara de la variabilidad normal dentro del conjunto.

- **Confirmación de hipótesis**

1) Tipo de moneda: ¿la fuerte dispersión salarial podría estar influenciada por la dolarización de parte de los sueldos?

Hallazgos

- Los sueldos en dólares están por encima de los sueldos en pesos, incluso después de la conversión.
- La mediana salarial de los que cobran el sueldo dolarizado es más alta que aquellos que cobran en pesos, y mayor rango intercuartílico. Quienes cobran en dólares están más de 1 millón arriba.
- La densidad de los salarios en pesos está más concentrada y en dólares más aplanada y extendida, indicando más heterogeneidad en los dolarizados. lo cual puede mostrar que puede haber distintos grados de dolarización o puestos de alta jerarquía.
- El grupo con sueldos dolarizados no solo tiene sueldos más altos (como seguramente viste en los promedios), sino que también es más desigual internamente. Esto puede deberse a que dentro del grupo dolarizado hay personas con cargos muy distintos: desde junior hasta senior o con distintas proporciones del sueldo en dólares.

Todo esto nos lleva a decir que nuestra hipótesis es real, la dolarización de los sueldos logra influenciar la dispersión salarial.

2) Modalidad de trabajo: ¿se presume que quienes trabajan de forma remota o híbrida tienen mayores ingresos?

Hallazgos:

- El trabajo remoto se asocia con mejores salarios que el presencial.
 - Los salarios en modalidad remota tienen mayor mediana y menor concentración de valores bajos. La dispersión es también mayor, lo que sugiere una brecha salarial más amplia en remoto.
 - La modalidad remota lidera en mediana, rango intercuartílico y máximos.
 - El trabajo híbrido se ubica en una posición intermedia, con características más similares al remoto que al presencial.
- Los hallazgos no llevan a decir que nuestra hipótesis también se comprueba, en el trabajo en remoto se gana más, y en segundo lugar el Híbrido, puede deberse a contratos internacionales, mayor flexibilidad o empresas extranjeras, o freelance. Quienes trabajan 100% remoto presentan un salario promedio bruto superior a quienes lo hacen presencial o en esquema híbrido.
- El trabajo presencial está asociado a menores ingresos y menor variabilidad.

3) Área tecnológica: ¿el rol profesional puede ser determinante en la remuneración?

Hallazgos:

- Muestra una fuerte dispersión salarial dentro de cada rol
- Permite observar que los roles de liderazgo y arquitectura no solo tienen mayores promedios salariales, sino también mayor variabilidad.
- Los puestos con mayor nivel de decisión o expertise técnico tienden a estar mejor remunerados

El rol profesional es un factor determinante en la remuneración en el sector IT. Los puestos de liderazgo y arquitecturas técnicas (Architect, Technical Leader, Manager) tienden a percibir salarios mayores que roles de soporte o analítica (HelpDesk, QA, BI Analyst). Además, se observa variabilidad dentro de los mismos roles, lo que también puede deberse a experiencia, empresa o región.