

INFORME

- **Título:** Análisis Profundo del Mercado Laboral en Ciencia de Datos: Un Enfoque Estadístico y Visual
- **Fecha de Elaboración:** 24/03/2024
- **Elaborado por:** Víctor Manuel Rocha Ramos

RESUMEN EJECUTIVO

El informe proporciona un análisis detallado del mercado laboral en ciencia de datos, utilizando datos extraídos de Kaggle. Se emplearon herramientas como Python, Pandas y MySQL para limpiar, analizar y extraer información de la base de datos. Además, se crearon visualizaciones dinámicas en Excel, Power BI y Tableau para facilitar la interpretación de los resultados.

Los hallazgos muestran una asociación entre el nivel de experiencia y los salarios, destacando roles como Ingeniero de Datos y Científico de Datos con los salarios más altos. El empleo a tiempo completo es el más común en la industria. Además, se observa una concentración geográfica en países como Estados Unidos, Reino Unido y Canadá.

Estos resultados ofrecen insights valiosos para profesionales y empresas en ciencia de datos, informando estrategias de contratación y desarrollo profesional.

INFORMACION GENERAL

1. **work_year:** El año en que se pagó el salario, lo que permite analizar las tendencias de compensación a lo largo del tiempo específicamente en la ciencia de datos.
2. **experience_level:** Clasifica a los profesionales según su nivel de experiencia en la ciencia de datos, dividido en:
 - **EN** (Entry level/Nivel de entrada): Profesionales con menos de 2 años de experiencia.
 - **MI** (Mid-level/Nivel medio): Aquellos con 2 a 5 años de experiencia.
 - **SE** (Senior level/Nivel senior): Profesionales con más de 5 años de experiencia.
 - **EX** (Expert/Experto): Reconocidos líderes en el campo con muchos años de experiencia y contribuciones significativas.
3. **employment_type:** Tipo de empleo del individuo en la ciencia de datos, incluyendo:
 - **FT** (Full-Time/Tiempo completo): Trabajo de horario completo estándar.
 - **PT** (Part-Time/Tiempo parcial): Empleo por menos horas que el completo.
 - **CT** (Contract/Contrato): Empleo bajo contrato por un periodo específico.
 - **FL** (Freelance/Freelancer): Profesionales independientes trabajando por proyectos.
4. **job_title:** El rol específico desempeñado dentro del ámbito de la ciencia de datos durante el año, reflejando la diversidad de especializaciones en este campo.
5. **salary:** Monto total del salario bruto pagado, fundamental para análisis de compensación.
6. **salary_currency:** La moneda en la que se pagó el salario, codificado según ISO 4217. Esto es relevante para realizar conversiones de moneda cuando se analizan salarios a nivel internacional.
7. **salaryinusd:** El salario convertido a dólares estadounidenses, lo que estandariza las comparaciones de compensación a través de diferentes países y economías.
8. **employee_residence:** Código ISO 3166 del país de residencia del empleado, crucial para estudios sobre cómo la ubicación geográfica afecta los salarios y las oportunidades en ciencia de datos.
9. **remote_ratio:** El porcentaje de trabajo realizado de manera remota, ofreciendo insights sobre la prevalencia y efectos del trabajo remoto en la ciencia de datos.
10. **company_location:** El país de la oficina principal del empleador, proporcionando una perspectiva sobre dónde están basadas las compañías que emplean profesionales de ciencia de datos.

11. **company_size**: El tamaño de la compañía, medido por el número de empleados, lo que puede influir en las estructuras de compensación y las oportunidades de carrera en la ciencia de datos.

Este resumen integra las dimensiones clave de empleo y compensación en el campo de la ciencia de datos, destacando cómo diferentes factores como el tipo de empleo, el nivel de experiencia, y la ubicación geográfica del empleado y del empleador pueden influir en los salarios y las oportunidades dentro de este sector en crecimiento.

INTRODUCCION

Contexto: En la era digital actual, la ciencia de datos ha emergido como un campo fundamental que impulsa decisiones estratégicas en diversas industrias. Desde la personalización de servicios hasta la optimización de procesos empresariales, el análisis de datos se ha convertido en un componente crucial para el éxito organizacional.

Fuente de Datos: La base de datos utilizada en este análisis fue extraída de Kaggle, una plataforma reconocida por albergar conjuntos de datos relevantes y de alta calidad. La selección de esta base de datos específica se justifica por su exhaustividad y representatividad en cuanto a la información sobre empleo en el campo de la ciencia de datos. Estos datos son fundamentales para ofrecer una perspectiva integral y confiable del mercado laboral en esta industria.

Objetivo del Análisis: El presente análisis tiene como objetivo principal profundizar en el mercado laboral de la ciencia de datos, abordando varios aspectos clave. Entre los objetivos específicos se incluye comprender la variación salarial por experiencia, identificar los tipos de empleo más prevalentes en la industria y explorar la distribución geográfica tanto de empleados como de empleadores.

ANALISIS DE DATOS Y RESULTADOS

PREGUNTAS OBJETIVO:

Pregunta 1: ¿Cómo varía el salario promedio por nivel de experiencia en ciencia de datos?

Respuesta:

El salario promedio varía significativamente según el nivel de experiencia en el campo de la ciencia de datos. Aquí están los salarios promedios por nivel de experiencia:

- Experto: \$194,931
- Nivel Senior: \$153,051
- Nivel Medio: \$104,526
- Nivel de Entrada: \$78,546

Pregunta 2: ¿Cuál es el tipo de empleo más común en la industria de la ciencia de datos?

Respuesta:

El tipo de empleo más común en la industria de la ciencia de datos es el tiempo completo (Full-Time), seguido por opciones mucho menos comunes como medio tiempo (Part-Time), contrato (Contract) y freelance. Los números exactos son:

- Tiempo completo (Full-Time): 3718
- Medio tiempo (Part-Time): 17
- Contrato (Contract): 10
- Freelance: 10

Pregunta 3: ¿Existe una diferencia significativa en el salario entre empleados a tiempo completo y empleados contratados?

Respuesta:

Sí, existe una diferencia significativa en el salario total acumulado entre diferentes tipos de empleo, siendo los empleados a tiempo completo los que acumulan la mayor cantidad:

- Tiempo completo (Full-Time): \$514,252,194
- Contrato (Contract): \$1,134,469
- Medio tiempo (Part-Time): \$672,073

- Freelance: \$518,078

Pregunta 4: ¿Qué nivel de experiencia tiende a estar asociado con mayores salarios en la ciencia de datos?

Respuesta:

Los niveles de experiencia que tienden a estar asociados con mayores salarios en la ciencia de datos son aquellos relacionados con ciertos títulos de trabajo. Aquí están los salarios totales asociados con algunos títulos de trabajo específicos:

- Ingeniero de Datos (Data Engineer): \$148,497,164
- Científico de Datos (Data Scientist): \$118,330,603
- Analista de Datos (Data Analyst): \$66,534,331
- Ingeniero de Aprendizaje Automático (Machine Learning Engineer): \$44,705,620
- Arquitecto de Datos (Data Architect): \$16,333,091

Esto sugiere que los roles como Ingeniero de Datos y Científico de Datos están asociados con los salarios más altos en la ciencia de datos.

Pregunta 5: ¿Cuál es la distribución de tamaños de empresa entre los empleadores de ciencia de datos?

Respuesta:

La distribución de tamaños de empresa entre los empleadores de ciencia de datos se inclina hacia las empresas medianas, seguidas por las grandes y luego las pequeñas. Los números son los siguientes:

- Empresa mediana: 3153
- Empresa grande: 454
- Empresa pequeña: 148

Pregunta 6: ¿Cómo difieren los salarios promedio entre pequeñas, medianas y grandes empresas en ciencia de datos?

Respuesta:

Los salarios promedio difieren entre pequeñas, medianas y grandes empresas, siendo las empresas medianas las que ofrecen los salarios promedio más altos:

- Empresa mediana: \$143,130.55

- Empresa grande: \$118,300.98
- Empresa pequeña: \$78,226.68

Pregunta 7: ¿Cuál es la ubicación geográfica más común para empleados y empleadores en la ciencia de datos?

Respuesta para empleadores (company_location):

- Estados Unidos (US): 3040
- Reino Unido (GB): 172
- Canadá (CA): 87
- España (ES): 77
- India (IN): 58

Respuesta para empleados (employee_residence):

- Estados Unidos (US): 3004
- Reino Unido (GB): 167
- Canadá (CA): 85
- España (ES): 80
- India (IN): 71

Pregunta 8: ¿Existe alguna relación entre el nivel de experiencia y el tipo de empleo ofrecido en la industria de la ciencia de datos?

Respuesta:

Sí, hay una relación entre el nivel de experiencia y el tipo de empleo ofrecido en la industria de la ciencia de datos. Se observa que ciertos niveles de experiencia tienden a ser más comunes en ciertos tipos de empleo. A continuación, se presenta la relación entre el nivel de experiencia y el tipo de empleo ofrecido:

- **Nivel de experiencia: Senior level**
 - Tiempo completo (Full-Time): 2511
 - Freelance: 3
- **Nivel de experiencia: Mid-level**
 - Tiempo completo (Full-Time): 792
 - Contrato (Contract): 5

- Freelance: 5
- Medio tiempo (Part-Time): 3
- **Nivel de experiencia: Entry level**
 - Tiempo completo (Full-Time): 302
 - Contrato (Contract): 2
 - Medio tiempo (Part-Time): 14
- **Nivel de experiencia: Expert**
 - Tiempo completo (Full-Time): 113

Estos datos muestran cómo diferentes niveles de experiencia están asociados con distintos tipos de empleo en la industria de la ciencia de datos. Por ejemplo, los roles de nivel senior son más comunes en empleos a tiempo completo, mientras que los roles de nivel de entrada se encuentran en una variedad de tipos de empleo, incluyendo tiempo completo, contrato y medio tiempo.

Pregunta 9: ¿Cuál es el salario promedio para diferentes tipos de empleo en empresas de diferentes tamaños?

Respuesta:

El salario promedio para diferentes tipos de empleo en empresas de diferentes tamaños se presenta a continuación:

- **Contrato (Contract)**
 - Empresa Grande (Large company): \$185,000.0
 - Empresa Mediana (Medium company): \$58,242.0
 - Empresa Pequeña (Small company): \$132,875.0
- **Freelance**
 - Empresa Grande (Large company): \$20,000.0
 - Empresa Mediana (Medium company): \$61,511.0
 - Empresa Pequeña (Small company): \$47,631.0
- **Tiempo completo (Full-Time)**
 - Empresa Grande (Large company): \$119,376.0
 - Empresa Mediana (Medium company): \$143,521.0
 - Empresa Pequeña (Small company): \$79,236.0

- **Medio tiempo (Part-Time)**

- Empresa Grande (Large company): \$32,724.0
- Empresa Mediana (Medium company): \$47,553.0
- Empresa Pequeña (Small company): \$39,661.0

Estos valores muestran las diferencias en los salarios promedio para distintos tipos de empleo, según el tamaño de la empresa.

Pregunta 10: ¿Qué países tienen la mayor concentración de empresas que contratan profesionales de ciencia de datos?

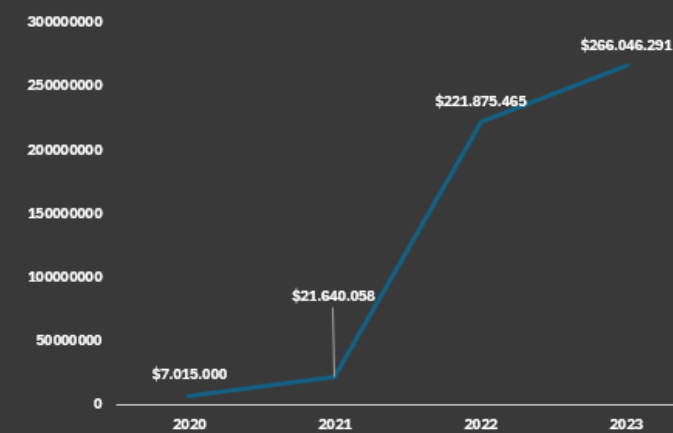
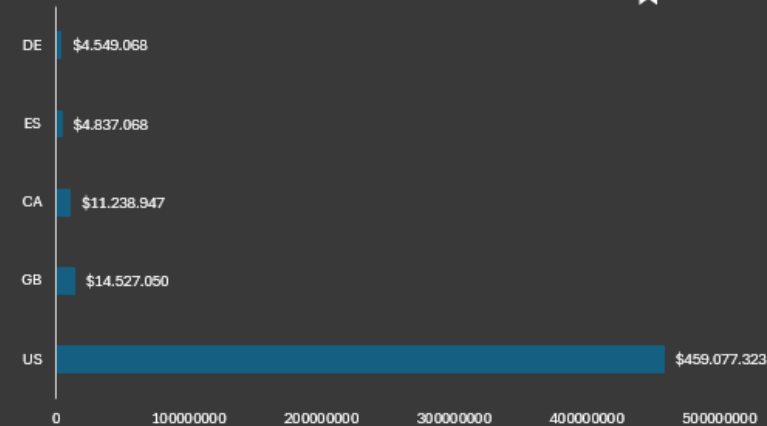
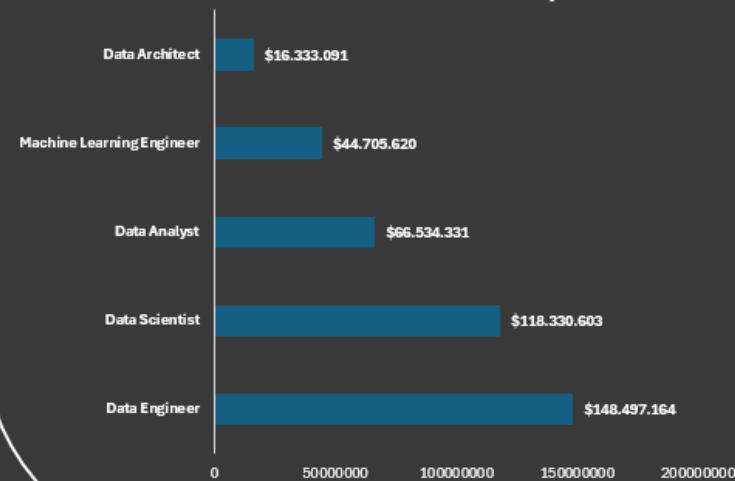
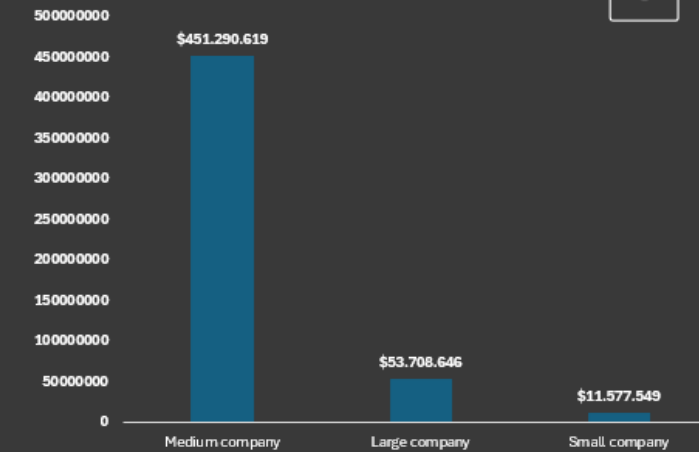
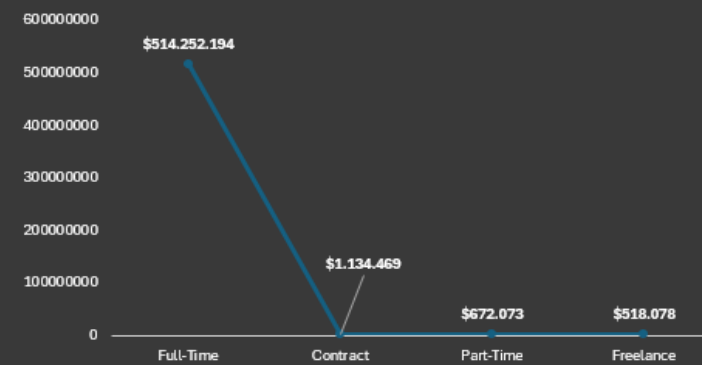
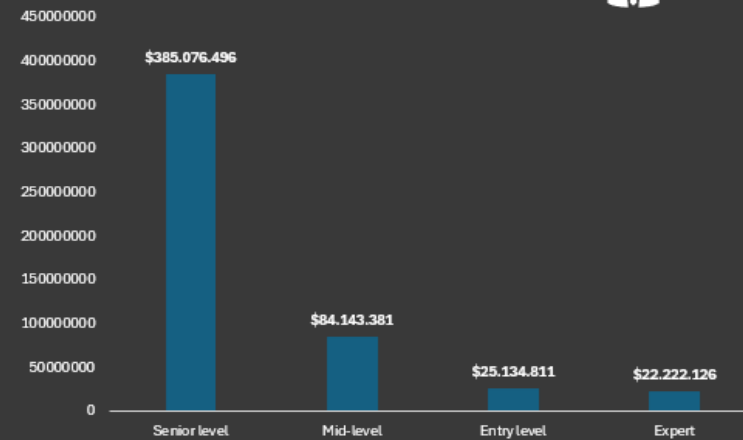
Respuesta:

La distribución de salarios para diferentes títulos de trabajo en la industria de la ciencia de datos se presenta a continuación:

- Ingeniero de Datos (Data Engineer): \$148,497,164
- Científico de Datos (Data Scientist): \$118,330,603
- Analista de Datos (Data Analyst): \$66,534,331
- Ingeniero de Aprendizaje Automático (Machine Learning Engineer): \$44,705,620
- Arquitecto de Datos (Data Architect): \$16,333,091

Estos números reflejan la variabilidad en los salarios asociados con diferentes roles en la industria de la ciencia de datos, con Ingeniero de Datos y Científico de Datos ocupando los rangos más altos.

DASHBOARD EXCEL



Año

2020

2021

2022

2023

Carrera

3D Computer Vision ...

AI Developer

AI Programmer

AI Scientist

Tipo Empleo

Contract

Freelance

Full-Time

Part-Time

Nivel Exp

Entry level

Expert

Mid-level

Senior level

Tamaño Compañía $\frac{1}{2}$ =
 $\frac{1}{3}$ =
 $\frac{1}{4}$ =

Large company

Medium company

Small company

DASHBOARD POWER BI

SALARIOS ANUAL DE DATA SCIENCE

FILTROS

Nivel Exp

- ☐ Entry level
- ☐ Expert
- ☐ Mid-level
- ☐ Senior level

Año

- 2020
- 2023



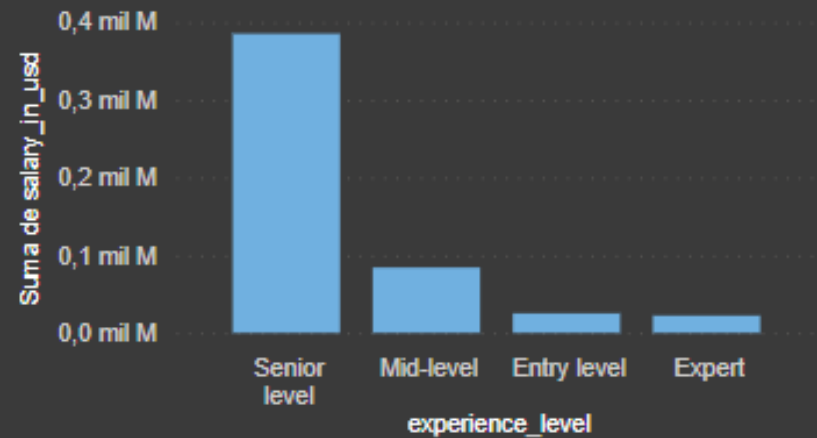
Tipo Empleo

- ☐ Contract
- ☐ Freelance
- ☐ Full-Time
- ☐ Part-Time

Tamaño Compañía

- ☐ Large company
- ☐ Medium company
- ☐ Small company

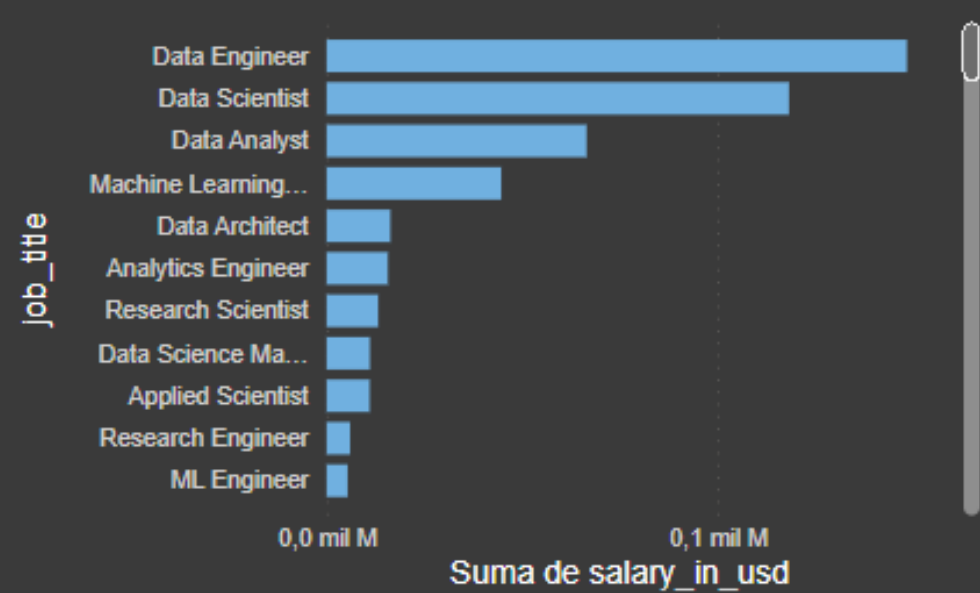
Salario por Nivel de Experiencia



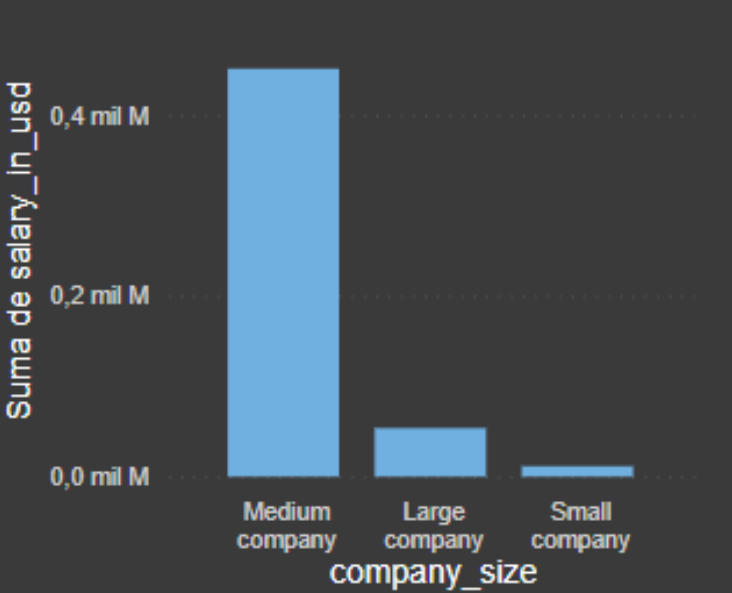
Salario por Tipo de Contrato



Salarios por Carrera

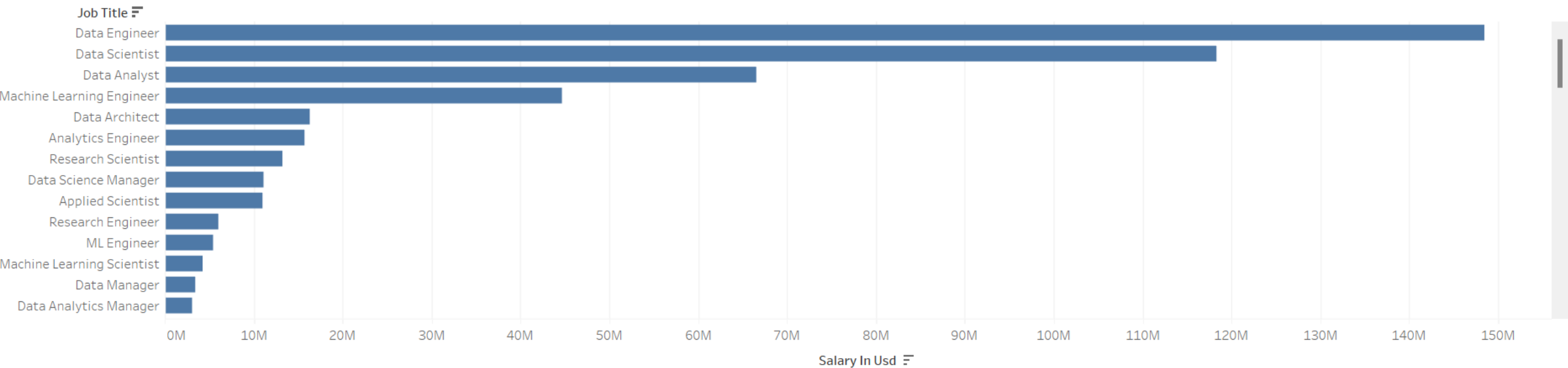


Salario por Tamaño de la Empresa

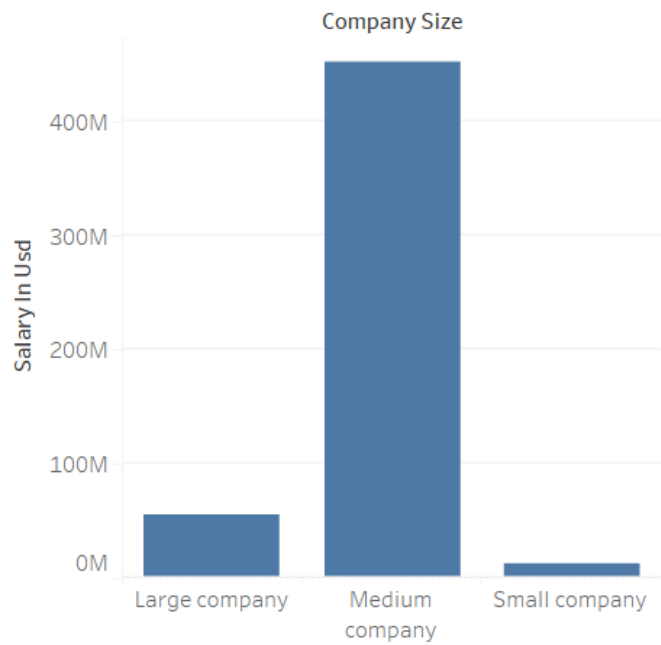


DASHBOARD TABLEAU

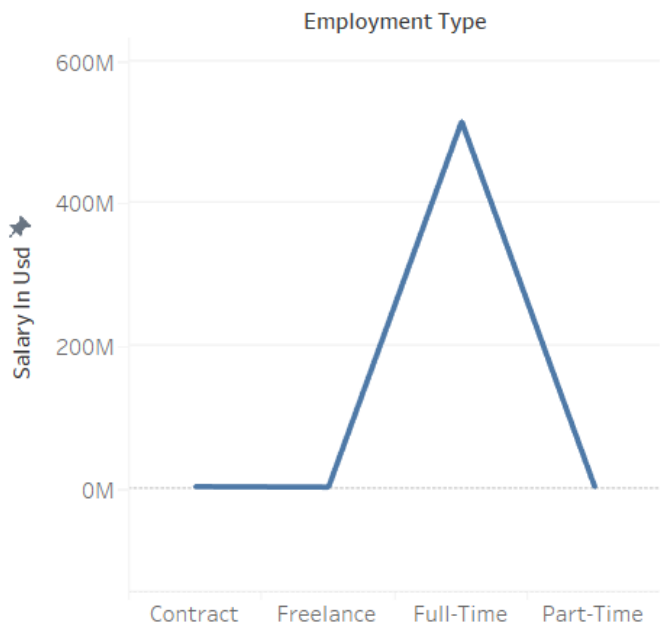
Salarios por Carrera



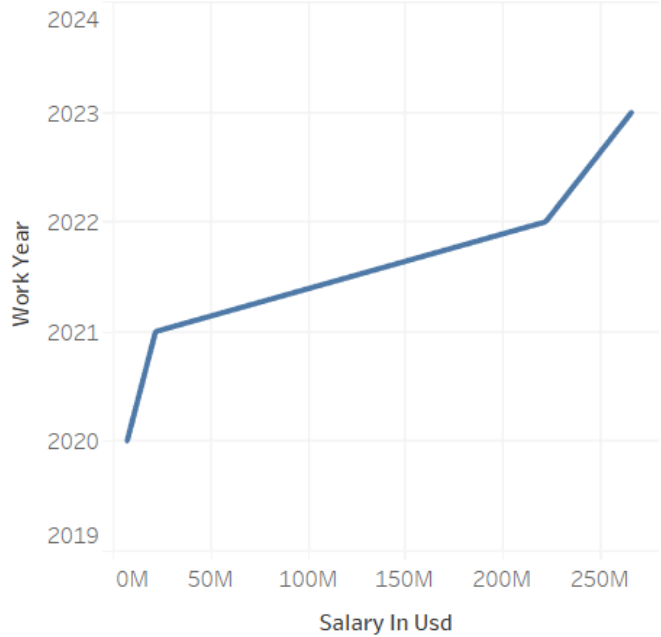
Salarios por Tamaño de Empresa



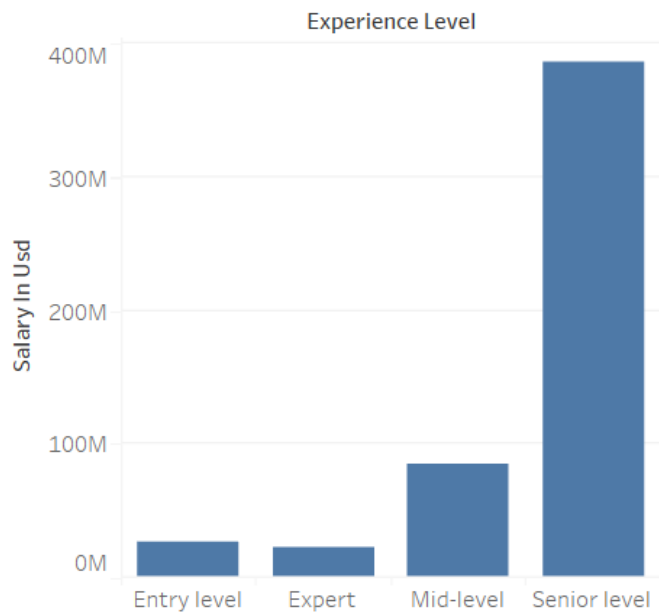
Salario por Tipo de Contrato



Salarios por Año



Salarios por Nivel de Experiencia



CONCLUSIONES

1. Variación Salarial por Experiencia:

- Se observa una tendencia clara de que los salarios aumentan con el nivel de experiencia en la industria de la ciencia de datos. Los roles de nivel senior tienden a tener los salarios más altos, seguidos de los roles de nivel medio, de experto y de entrada.

2. Tipos de Empleo Más Comunes:

- El empleo a tiempo completo es el más común en la industria de la ciencia de datos, seguido por empleos a tiempo parcial, contratos y trabajo freelance. Esto refleja una demanda predominante de empleados a tiempo completo en el campo.

3. Diferencias Salariales según el Tamaño de la Empresa:

- Existen diferencias significativas en los salarios promedio entre empresas de diferentes tamaños. Las empresas medianas ofrecen los salarios promedio más altos, seguidas de las grandes empresas y luego las pequeñas empresas.

4. Ubicaciones Geográficas Relevantes:

- Estados Unidos (US) es la ubicación geográfica más común tanto para empleados como para empleadores en la industria de la ciencia de datos, seguido de Reino Unido (GB) y Canadá (CA).

5. Relación entre Experiencia y Tipo de Empleo:

- Se observa una asociación entre el nivel de experiencia y el tipo de empleo ofrecido. Los roles de nivel senior tienden a ser más comunes en empleos a tiempo completo, mientras que los roles de nivel de entrada se encuentran en una variedad de tipos de empleo, incluyendo tiempo completo, contrato y medio tiempo.

Estas conclusiones destacan las tendencias clave en el mercado laboral de la ciencia de datos, proporcionando información valiosa para empresas, profesionales y tomadores de decisiones en este campo en constante evolución.

REFERENCIAS:

Fuente de datos extraída de: <https://www.kaggle.com/datasets/arnabchaki/data-science-salaries-2023>