



---

**PROGRAMACIÓN PARA ANÁLISIS DE DATOS - PROYECTO EN R**

---

**Autores:**

Renzo Giacometti - 5.590.459-8

Francisco Nasso - 5.381.393-5

**Año 2025**

# Introducción

## **Dataset elegido**

Para este proyecto se seleccionó el dataset titulado “**AI Impact on Job Market: Increasing vs Decreasing Jobs (2024–2030)**”, disponible públicamente en Kaggle. El archivo contiene información sintética pero realista basada en estudios de organismos como U.S. BLS, OECD, McKinsey y el WEF.

<https://www.kaggle.com/datasets/sahilislam007/ai-impact-on-job-market-20242030>

El dataset cuenta con 30.000 filas y 13 variables que describen distintos aspectos del mercado laboral, incluyendo el nivel de impacto de la IA sobre los puestos de trabajo, el riesgo de automatización, salarios, educación requerida, diversidad de género y otras características relevantes.

## **Pregunta a responder y fundamentación**

### **Pregunta principal:**

¿Cómo está afectando la adopción de la Inteligencia Artificial al mercado laboral global en el período 2024–2030?

### **Fundamentación:**

Esta pregunta resulta clave en un contexto donde la transformación tecnológica avanza rápidamente, modificando no solo las herramientas de trabajo, sino también la estructura misma del empleo. A través de este análisis, buscamos identificar qué tipo de trabajos están creciendo o desapareciendo, qué sectores se ven más afectados, y cómo variables como el salario, el riesgo de automatización o el nivel educativo influyen en estas dinámicas.

Este análisis puede resultar útil para:

- **Gobiernos:** en el diseño de políticas públicas de formación y reconversión laboral
- **Empresas:** para planificar estrategias de contratación y automatización.
- **Trabajadores y estudiantes:** para orientar sus decisiones vocacionales y formativas.

- **Investigadores:** para explorar correlaciones y efectos indirectos de la IA sobre el empleo.

## Análisis de datos

### Estructura general

- **Filas:** 30.000
- **Columnas:** 13

### ¿Qué representa cada fila?

Cada fila del dataset representa un puesto de trabajo específico en determinado país, con información proyectada sobre su evolución entre 2024 y 2030 en el contexto de la adopción creciente de tecnologías basadas en inteligencia artificial.

### Tipos de variables

- **Categorías:**
  - Job Title
  - Industry
  - Job Status
  - AI Impact Level
  - Required Education
  - Location
- **Numéricas:**
  - Median Salary (USD)
  - Experience Required (Years)
  - Job Openings (2024)
  - Projected Openings (2030)
  - Remote Work Ratio (%)
  - Automation Risk (%)

- Gender Diversity (%)

### **Análisis de duplicados y de valores faltantes**

Se examinó la presencia de registros duplicados completos. En este caso no se encontró ningún duplicado, por lo que no se tuvo que realizar nada al respecto.

En cuanto a los valores faltantes, tampoco se identificó ninguno dentro del dataset, por lo que se prosiguió sin problema.

### **Análisis exploratorio de datos**

A continuación se presentan visualizaciones y cálculos descriptivos que permiten responder cómo la IA está afectando al mercado laboral en el período 2024–2030.

#### **Estadísticas Descriptivas**

Decidimos tomar datos respecto al salario de los empleos afectados por IA, ya que al observar los salarios se puede intentar trazar una imagen de qué estilo de trabajos tenderán a ser automatizados por esta herramienta. También lo identificamos como el dato del que más información se podría tomar con este análisis.

Median Salary USD:

Min: 30002

1st Qu: 60501

Median: 90274

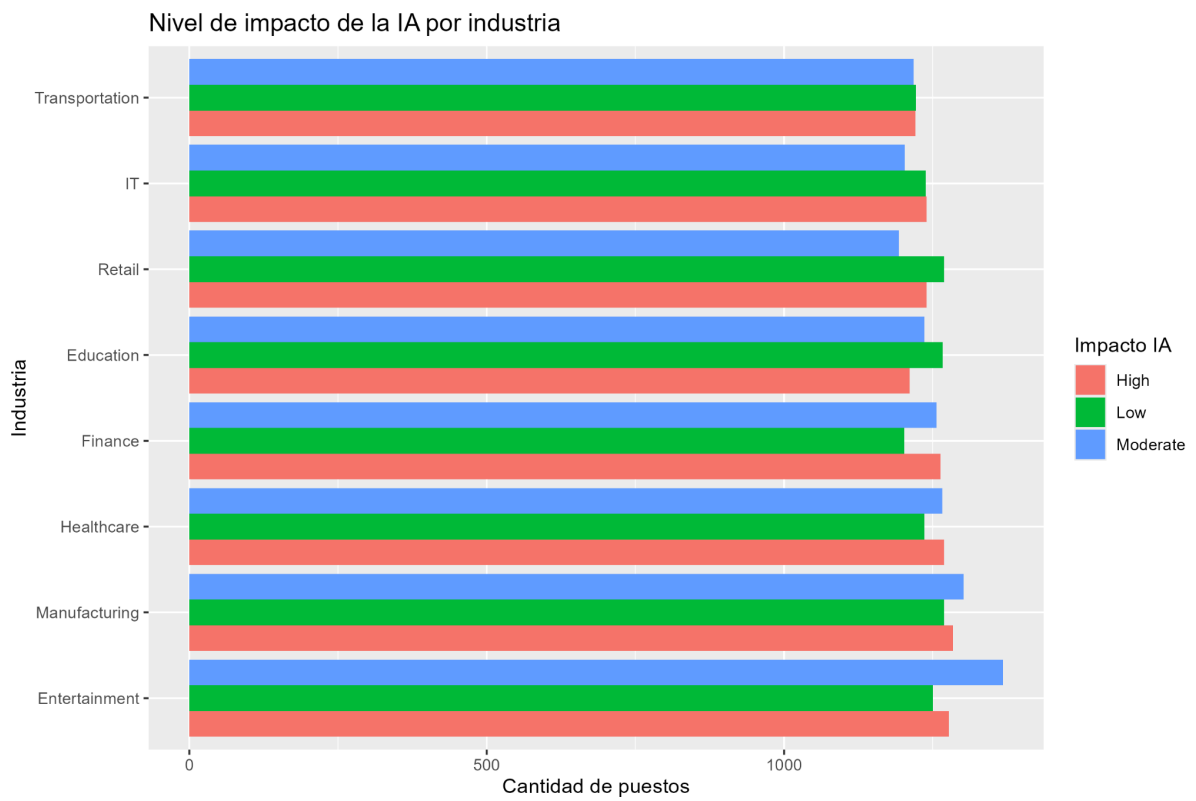
Mean 3rd Qu: 119455

Max: 149999

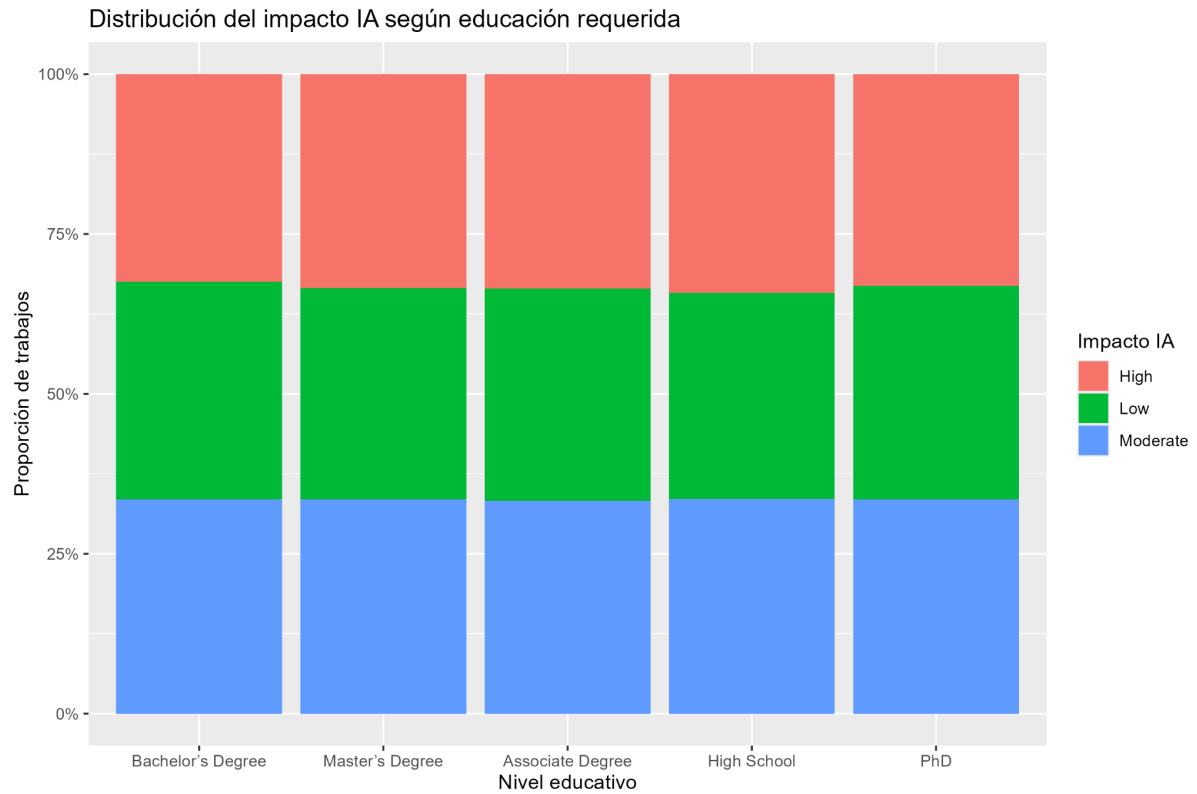
---

## Gráficos con ggplot2

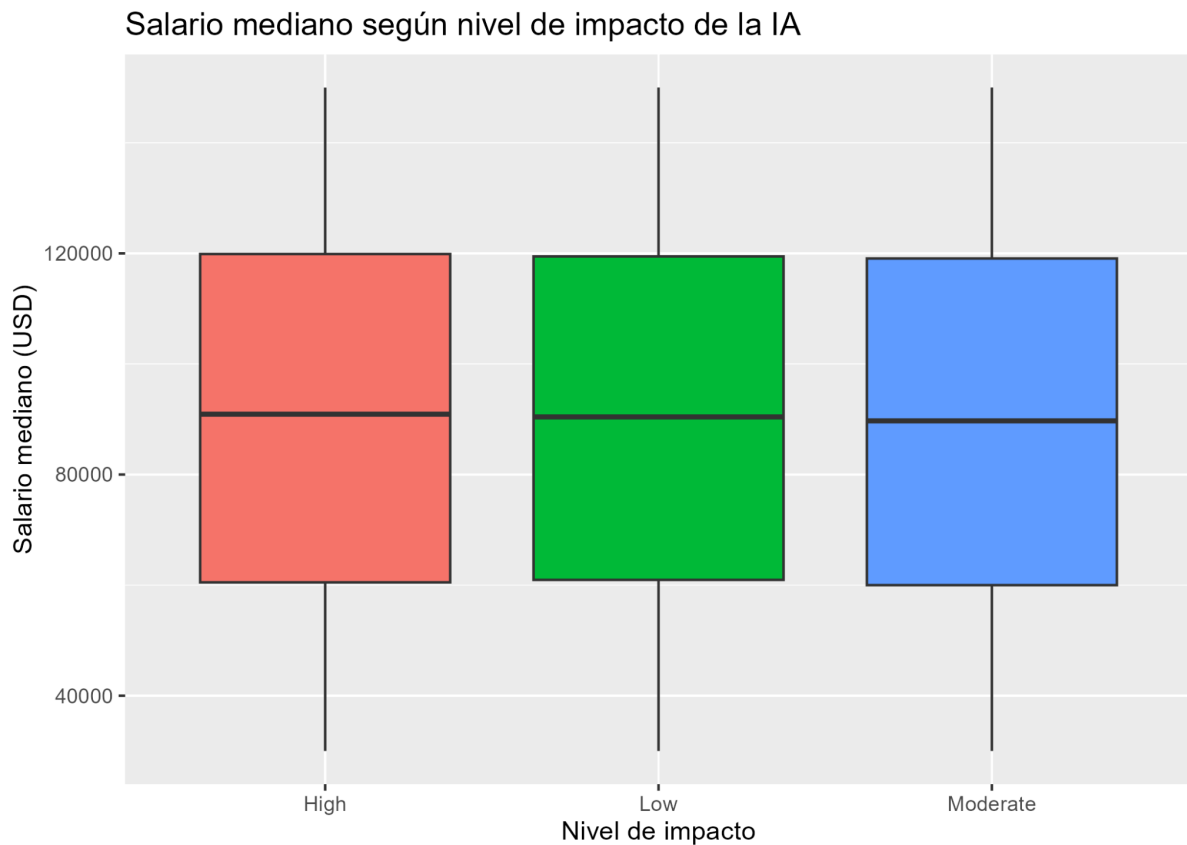
### 1. Gráfico de barras – Nivel de impacto de IA por industria de trabajo



## 2. Gráfico de barras apiladas – Impacto de IA por nivel educativo



### 3. Boxplot – Salario por Nivel de impacto



## Conclusiones

Frente a las hipótesis que se pudieran tener al iniciar este análisis, podemos observar que no existe una gran diferencia de afectación de IA entre las distintas industrias o entre distintos rangos salariales. Al mirar de cerca la información del dataset, nos parece importante resaltar que es un estudio enfocado en una predicción para 2030, tiempo que puede resultar no suficiente para la adopción de esta tecnología al nivel en que afecte en gran medida a una industria entera. Pese a que se están observando avances, cargos elevados, manuales, y más artesanales o sociales dentro de todas las industrias siguen siendo difíciles de reemplazar, y existen todos los ámbitos laborales, por lo que parece que aún habrá que esperar un tiempo más

largo para poder notar una adopción en gran escala que tome en gran medida una industria, o un rango salarial.