

# ANLY500 Project

Abhishek Anand, Bolun Lu, Siming Xu, Rijin Lnu

2022-10-02

## **Salaries in the Fields of Data Science, Artificial Intelligence, and Machine Learning**

### **Introduction**

A shift towards data-centered decision-making is taking hold and at the center of this shift are those working in the fields of data science (“DS”), artificial intelligence (“AI”), and machine learning (“ML”). As businesses seek to optimize operations and information using vastly increased computing power, the shape and size of data analytics has increased exponentially over the last two decades, and even more prominently in the last decade with the rise of machine learning, cloud computing, and advances in the field of artificial intelligence.

Exploring the composition and trends of salaries in this sector is equally as intriguing, both from a research perspective and from the perspective of what can be waiting for students of Data Analytics (DS/AI/ML) in academic programs globally.

The data we’ve chosen for this project comes from <https://salaries.ai-jobs.net/>, a website collecting salary data from around the world.

This data set has salary (in USD) for a variety of technical roles, and a variety of information on the characteristics of the employee, employer, the job itself, and the market.

The data set contains the following variables:

<b>work_year</b>	The year the salary was paid.								
<b>experience_level</b>	<p>The experience level in the job during the year with the following possible values:</p> <table> <tr> <td><b>EN</b></td><td>Entry-level / Junior</td></tr> <tr> <td><b>MI</b></td><td>Mid-level / Intermediate</td></tr> <tr> <td><b>SE</b></td><td>Senior-level / Expert</td></tr> <tr> <td><b>EX</b></td><td>Executive-level / Director</td></tr> </table>	<b>EN</b>	Entry-level / Junior	<b>MI</b>	Mid-level / Intermediate	<b>SE</b>	Senior-level / Expert	<b>EX</b>	Executive-level / Director
<b>EN</b>	Entry-level / Junior								
<b>MI</b>	Mid-level / Intermediate								
<b>SE</b>	Senior-level / Expert								
<b>EX</b>	Executive-level / Director								
<b>employment_type</b>	<p>The type of employment for the role:</p> <table> <tr> <td><b>PT</b></td><td>Part-time</td></tr> <tr> <td><b>FT</b></td><td>Full-time</td></tr> <tr> <td><b>CT</b></td><td>Contract</td></tr> <tr> <td><b>FL</b></td><td>Freelance</td></tr> </table>	<b>PT</b>	Part-time	<b>FT</b>	Full-time	<b>CT</b>	Contract	<b>FL</b>	Freelance
<b>PT</b>	Part-time								
<b>FT</b>	Full-time								
<b>CT</b>	Contract								
<b>FL</b>	Freelance								
<b>job_title</b>	The role worked in during the year.								
<b>salary</b>	The total gross salary amount paid.								
<b>salary_currency</b>	The currency of the salary paid as an ISO 4217 currency code.								
<b>salary_in_usd</b>	The salary in USD (FX rate divided by avg. USD rate for the respective year via <a href="https://fxdata.foorilla.com">fxdata.foorilla.com</a> ).								
<b>employee_residence</b>	Employee's primary country of residence in during the work year as an ISO 3166 country code.								
<b>remote_ratio</b>	<p>The overall amount of work done remotely, possible values are as follows:</p> <table> <tr> <td><b>0</b></td><td>No remote work (less than 20%)</td></tr> <tr> <td><b>50</b></td><td>Partially remote</td></tr> <tr> <td><b>100</b></td><td>Fully remote (more than 80%)</td></tr> </table>	<b>0</b>	No remote work (less than 20%)	<b>50</b>	Partially remote	<b>100</b>	Fully remote (more than 80%)		
<b>0</b>	No remote work (less than 20%)								
<b>50</b>	Partially remote								
<b>100</b>	Fully remote (more than 80%)								
<b>company_location</b>	The country of the employer's main office or contracting branch as an ISO 3166 country code.								
<b>company_size</b>	<p>The average number of people that worked for the company during the year:</p> <table> <tr> <td><b>S</b></td><td>less than 50 employees (small)</td></tr> <tr> <td><b>M</b></td><td>50 to 250 employees (medium)</td></tr> <tr> <td><b>L</b></td><td>more than 250 employees (large)</td></tr> </table>	<b>S</b>	less than 50 employees (small)	<b>M</b>	50 to 250 employees (medium)	<b>L</b>	more than 250 employees (large)		
<b>S</b>	less than 50 employees (small)								
<b>M</b>	50 to 250 employees (medium)								
<b>L</b>	more than 250 employees (large)								

Figure 1: A caption