|  |  |  |  |
| --- | --- | --- | --- |
|  | Ariel Jones, Joe Liang, Kristina Mulholland, Jeff Pinegar | | jeffpinegar1@gmail.com |
| Assignment: | Project 3: Proposal | 717-982-0516 |
| Due Date: | Jan. 30, 2023 |  |

**Data Science Salaries**

# How much do people working in Data Science earn?

## Scope

* Data Science Salaries 2019 - 2022
* Data Source: Kaggle – Data Science Job Salaries  
  <https://www.kaggle.com/datasets/ruchi798/data-science-job-salaries?select=ds_salaries.csv>

## Questions

* Was the trend over the past 3 years?
* How big of a factor is experience or the size of the company where you work?
* How might your salary change with experience?
* How do salaries vary by job, location, experience, company size, etc.?
* How will working remotely affect my salary?

## Data Visuals

* Data Overview
* Average Salary by Location Analysis
* Salary Variations by Company Size
* Analysis by Experience
* Position’s Impact on Salary
* Salary Trends
* Machine Learning Dashboard

# Data Overview *(potential visuals and interactions)*

### Descriptive Summary Statistic for the variable in the data set.

* Job categories
* Remote work
* Employment Type
* Year
* Level of Experience
* Salary (Average, Median, Std. Dev.)

# Average Salary by Location Analysis *(potential visuals and interactions)*

|  |  |
| --- | --- |
| Graphic  * Map * Bar Chart * ? | Interactivity  * Year * Company Size * Employment Type * Job Category * Experience |

# Salary Variations by Company Size *(potential visuals and interactions)*

|  |  |
| --- | --- |
| Avg. Salary by Company Size and year  * Bar Chart * ? | Interactivity  * Country * Category * Employment Type * Experience |
| % of position by Category  * Bar Chart – the % of each category for each company size. | Interactivity  * Country * Employment Type * Experience level * Remote Work |
| Level of Experience with Company  * Bar Chart – Count of each category for each company size. | Interactivity  * Country * Employment Type * Category * Remote Work |

# Analysis by Experience *(potential visuals and interactions)*

|  |  |
| --- | --- |
| Experience Level’s Impact on Salary  * Bar Chart – Average salary | Interactivity  * Year * Country * Category * Company Size * Employment Type * Remote Work |

# Position’s Impact on Salary *(potential visuals and interactions)*

|  |  |
| --- | --- |
| Average Salary by Category  * Bar Chart | Interactivity  * Year * Country * Company Size * Employment Type * Experience |
| Salary by Job Title (Top -10)  * Bar Chart | Interactivity  * Year * Country * Company Size * Employment Type * Experience |
| Experience and Category  * Heat Map | Interactivity  * Year * Country * Company Size * Employment Type * Experience |
| Salary Distribution  * Histogram | Interactivity  * Year * Country * Company Size * Employment Type * Experience |

# Salary Trends *(potential visuals and interactions)*

|  |  |
| --- | --- |
| Experience Level’s Impact on Salary  * Bar Chart or line chart – Average salary by year | Interactivity  * Year * Country * Category * Company Size * Employment Type * Remote Work |

# Machine Learning (ML) *(potential visuals and interactions)*

|  |  |
| --- | --- |
| ML Prof vs None ML Prof.  * Line chart – trend over time | Interactivity  * Country * Company Size * Employment Type * Experience |
| Mix of experience in ML  * Bar Chart | Interactivity  * Year * Country * Company Size * Employment Type * Experience |
| ML representation by company size  * Heat Map | Interactivity  * Year * Country * Company Size * Employment Type * Experience |
| Average Salary by Position  * Histogram | Interactivity  * Year * Country * Company Size * Employment Type * Experience |

# Data Source

Data Source: Kaggle – Data Science Job Salaries  
<https://www.kaggle.com/datasets/ruchi798/data-science-job-salaries?select=ds_salaries.csv>

The data set includes these parameters:

|  |  |  |
| --- | --- | --- |
| Column | Description | Unique Values |
| work\_year | The year the salary was paid. | 3 |
| experience\_level | The experience level in the job during the year with the following possible values: EN Entry-level / Junior MI Mid-level / Intermediate SE Senior-level / Expert EX Executive-level / Director | 4 |
| employment\_type | The type of employment for the role: PT Part-time FT Full-time CT Contract FL Freelance | 4 |
| job\_title | The role worked in during the year. | 50 |
| salary | The total gross salary amount paid. | n.a. |
| salary\_currency | The currency of the salary paid as an ISO 4217 currency code. | 17 |
| salary\_in\_usd | The salary in USD (FX rate divided by avg. USD rate for the respective year via fxdata.foorilla.com). | n.a. |
| employee\_residence | Employee's primary country of residence in during the work year as an ISO 3166 country code. | 50 |
| remote\_ratio | The overall amount of work done remotely; possible values are as follows: 0 No remote work (less than 20%) 50 Partially remote 100 Fully remote (more than 80%) | 3 |
| company\_location | The country of the employer's main office or contracting branch as an ISO 3166 country code. | 50 |
| company\_size | The average number of people that worked for the company during the year: S less than 50 employees (small) M 50 to 250 employees (medium) L more than 250 employees (large) | 3 |