

REPORT - DATA SCIENCE CAREER ATLAS

Overview

This report presents an in-depth examination of a sophisticated Tableau dashboard that was meticulously designed to simplify the diverse data environment of global data science related occupations, including wage trends, employment categories, and regional compensation distributions. This dashboard, which draws on Kaggle's repositories, provides stakeholders with a helpful insight into the changing field of data science careers by including a broad temporal range from 2020 to 2023.

Data Handling and Visualization Tools

I used Tableau for this project because it's great for interactive stuff and really lets you get into the data. The datasets, meticulously curated from Kaggle, encompass a broad spectrum of occupational metrics within the realms of Artificial Intelligence, Machine Learning, and Data Science. They were pretty hefty, so I had to do some cleanup—lining them up by dates, kicking out any empty spots, and making sure similar data points played nice together when I mashed them into one big dataset.

Data Complexity and Analytical Scope

The datasets present a rich tapestry of attributes, blending quantitative aspects such as remuneration and chronological benchmarks, with qualitative and ordinal facets like employment nature and professional seniority. Such a dataset, characterized by its intricate high-dimensionality, necessitates an adroit visualization strategy to deftly unravel its complexities and distils actionable insights.

Visualization Strategies and Rationale

The dashboard underpins several analytical endeavours:

- **Comparative Analysis:** Comparing median salaries across different experience levels and employment types.
- **Temporal Trend Analysis:** Observing salary trends over time to gauge the growth or decline of earnings across employment types.
- **Geographical Distribution:** Mapping median salaries globally to identify regional disparities.
- **Correlational Analysis:** Examining the relationship between remote work trends and company size with compensation levels.

Encoding Channels and Idioms

The dashboard employs a mix of encoding channels and visualization idioms, each chosen for its efficacy in communicating specific types of data:

- **Geographical Maps:** Colours represent national median earnings, making cross-country comparisons easy to see.
- **Area charts:** Used to show pay trajectories over time. The employment type is indicated by the colour and size of the area, which provides a temporal salary progression.
- **Bar charts:** Allow for the comparison of median pay for various job roles across a range of experience levels. Experience levels are indicated by colour.
- **Bubble Charts:** These visual aids display the proportional demand for job titles. The size of the bubbles indicates the number of jobs, and the colour indicates the year.
- **Treemaps:** Make effective use of available space to show hierarchical data, like the size of the company compared to the type of employment, with the median wage indicated by the colour intensity.
- **Donut charts:** Using pieces of varying sizes, provide a clear visual description of the distribution of remote work.

Interactive operations include filter actions where selection on one visualization, like the map, dynamically updates others, fostering an exploratory data analysis experience.

Innovation and Implementation

This visualization's innovative feature is its integrative methodology. By combining temporal and spatial analysis, it surpasses conventional representations and provides a thorough overview uncommon in singledimensional charts. In order to maintain the dashboard's usability and educational value, it is necessary to strike a balance between the amount of information and its clarity and interactivity. This makes the implementation process hard. The synchronization of many data types and the synchronization of interactivity across dissimilar visual elements raise implementation complexity. With the use of interactive features and graphical elements that are easy to understand, the complexity of the presentation is controlled, making even rich datasets accessible.

Pros and Cons

The Good Stuff:

- **Interactivity:** Promotes an inquisitive and user-driven experience by allowing the user to explore deeper into particular areas of the data.
- **Extensive:** Provides a wide-ranging, yet intricate perspective on the field of data science and its associated job market.
- **Aesthetic Balance:** Preserves detail without compromising an eye-catching visual composition.
- Diverse needs and preferences for data analysis are accommodated by a wide variety of visualization types.
- The dashboard's design is unified, offering a user-friendly interface with a dependable colour palette and arrangement.
- Maps and bubble charts that employ colour and size effectively convey the data itself spread through a quick visual cue.

Could Be Better:

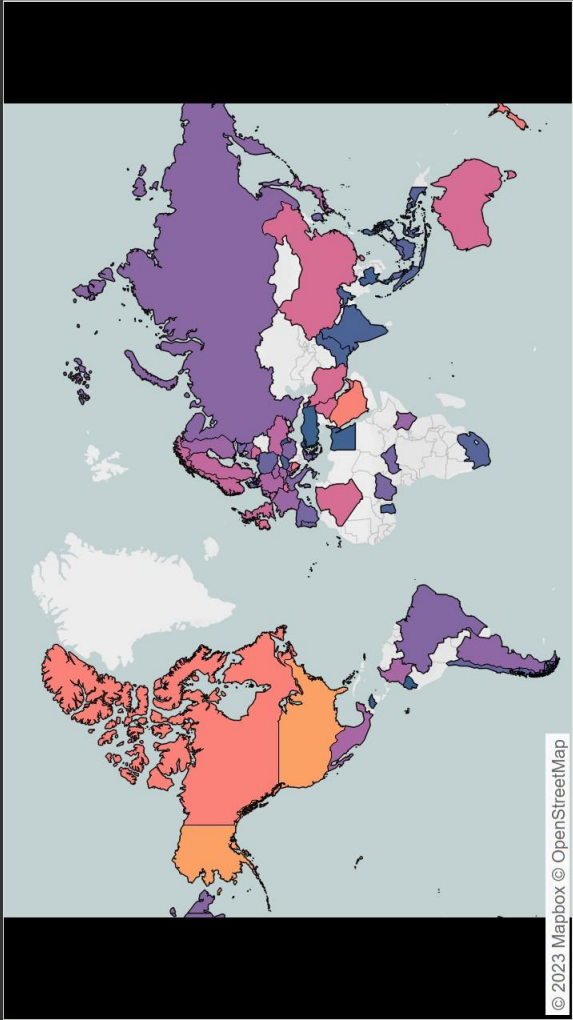
- **Information Density:** Some users may find the volume of information offered to be too much to handle, and there may be a learning curve.
- **Potential Bias:** Relying too much on median figures could make outliers harder to spot and distort perceptions of pay distributions.
- The analysis may be oversimplified when the distribution's range is obscured by the use of median statistics.
- Those who suffer from deficits in their colour vision may find it more difficult to rely on colour as their primary encoding channel.

Conclusion

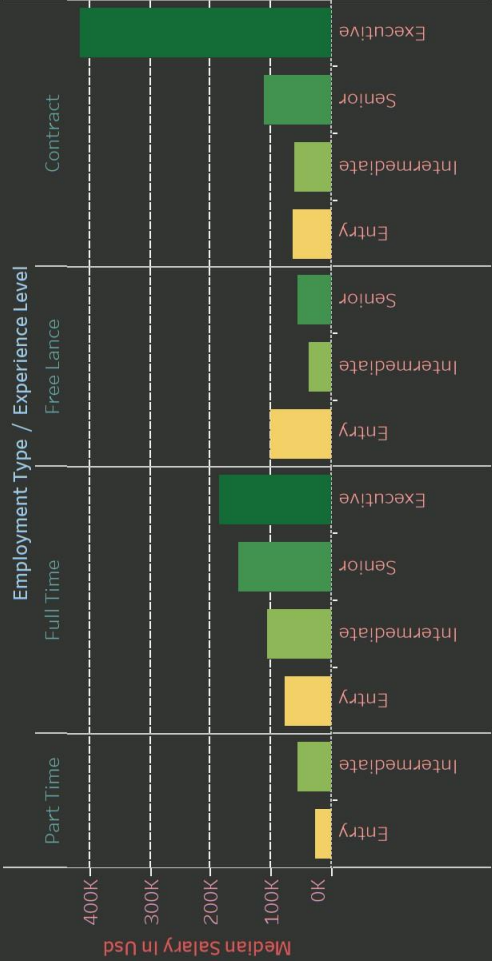
In addition to being a useful visualization tool, the "Data Science Career Atlas" skillfully breaks down the intricate relationships between the various data science employment markets. It sheds light on pay patterns, popular jobs, and preferences for remote work, giving stakeholders a comprehensive understanding of how the sector is changing. The dashboard provides a detailed analytical story, but how well the user navigates its complex visual encodings determines how effective the dashboard is. This highlights the need of intuitive design in data interpretation.

The Data Science Career Atlas: Global Earnings and Employment Dynamics

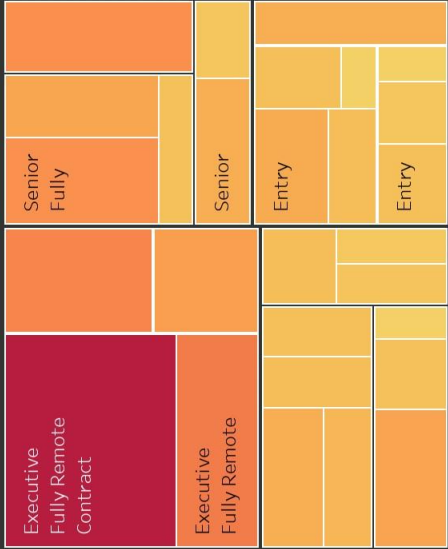
Global Distribution of Median Data Science Salaries by Country



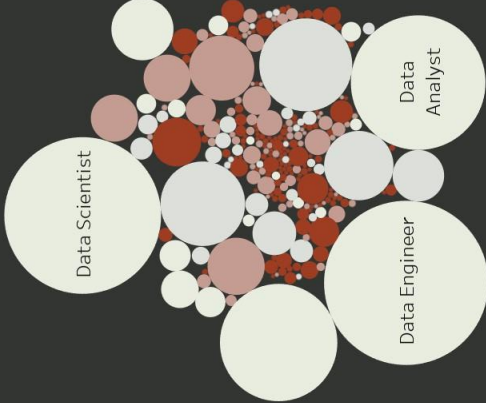
Earning Patterns by Role Seniority in Data Science Careers



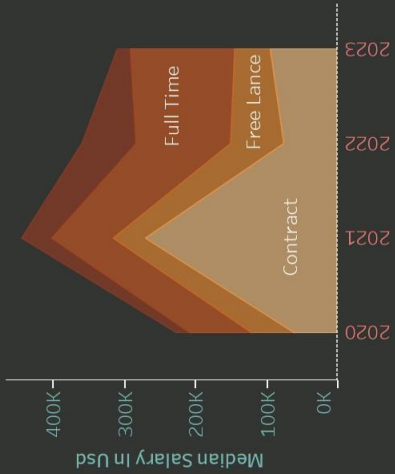
Corporate Ladder Salaries: A Data Science Industry Perspective



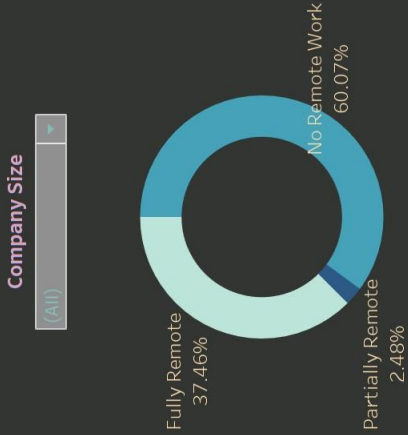
Popular Jobs in Data Science



Employment Type Salary Trajectories



Ring of Remote Work: Current Trends



REFERENCES:

1. Kaggle dataset - Global AI, ML, Data Science Salary 2023([link](#))
2. Kaggle dataset - Data Science Salary 2021 to 2023([link](#))
3. Blackboard notes
4. Tableau documentation([link](#))
5. Video Link – [Click here](#)(the video is present in the google drive, set to viewable by anyone). If the link above is not visible, copy and paste this URL on a browser –
https://drive.google.com/file/d/1B3TMMrSfdJ4AZ3gDD85DhtwOi8N06gE-/view?usp=drive_link

Note:

If the charts are unable to sense the data fields after opening the Tableau file, please right click on the database present on the data pane and click refresh. Even if doing that did not reflect in the charts, head over to the “Data source” pane present in the left end of Tableau and see if the check box – “Cleaned with data interpreter” to be ticked if not done. Even if it is done uncheck and check the box again and your issues might be solved and dashboard will be loaded.