Your CEO has decided that the company needs a full-time data scientist, and possibly a team of them in the future. She thinks she needs someone who can help drive data science within then entire organization and could potentially lead a team in the future. She understands that data scientist salaries vary widely across the world and is unsure what to pay them. To complicate matters, salaries are going up due to the great recession and the market is highly competitive. Your CEO has asked you to prepare an analysis on data science salaries and provide them with a range to be competitive and get top talent. The position can work offshore, but the CEO would like to know what the difference is for a person working in the United States. Your company is currently a small company but is expanding rapidly.

Prepare your analysis in an R file. Your final product should be a power point presentation giving your recommendation to the CEO. CEOs do not care about your code and don’t want to see it. They want to see visuals and a well thought out analysis. You will need to turn in the power point and the code as a flat R file.

The metadata are as follows:

| Column | Description |
| --- | --- |
| work\_year | The year the salary was paid. |
| 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 |
| employment\_type | The type of employement for the role: PT Part-time FT Full-time CT Contract FL Freelance |
| job\_title | The role worked in during the year. |
| salary | The total gross salary amount paid. |
| salary\_currency | The currency of the salary paid as an ISO 4217 currency code. |
| salary*in*usd | The salary in USD (FX rate divided by avg. USD rate for the respective year via fxdata.foorilla.com). |
| employee\_residence | Employee's primary country of residence in during the work year as an ISO 3166 country code. |
| 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%) |
| company\_location | The country of the employer's main office or contracting branch as an ISO 3166 country code. |
| 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) |

|  |  |
| --- | --- |
| Job Title | Description |
| 3D Computer Vision Engineer | Contribute to new features and 3D reconstruction strategies of the proprietary engine and associated applications. |
| AI Scientist | Experts in building algorithms. Responsible for creating new algorithms that can solve problems in new ways or improve existing ones. Usually work with data scientists and computer programmers to develop accurate and reliable models. |
| Analytics Engineer | Provide clean data sets to end users, modeling data in a way that empowers end users to answer their own questions. Spends more time transforming, testing, deploying and documenting data rather than analyzing data. |
| Applied Data Scientist | Analyze complex data sets, develop statistical models and apply machine learning techniques. |
| Applied Machine Learning Scientist | Execute the entire model development lifecycle, from prototype to delivery, while applying best practices for reproducible research and well-managed software delivery. |
| BI Data Analyst | Reviews data to produce finance and market intelligence reports. These reports are used to highlight patterns and trends in given market that may influence a company’s operations and future goals. |
| Big Data Architect | Expert who formulates the organizational data strategy, including standards of data quality, the flow of data within the organization, and security of data. |
| Big Data Engineer | Information technology (IT) professional who is responsible for designing, building, testing and maintaining complex data processing systems that work with large data sets. |
| Business Data Analyst | Use the data analysis process to advance their company’s business goals. Use data analysis process to understand, interpret, and predict patterns in business and then those data-driven insights to enhance their business practices. |
| Cloud Data Engineer | An IT professional responsible for any technological duties associated with cloud computing, including design, planning, management, maintenance, and support. |
| Computer Vision Engineer | Applies computer vision and machine learning research to solve real-world problems. Uses large sums of data and statistics in order to complete complex tasks and apply supervised or unsupervised learning as part of computer vision tasks. |
| Computer Vision Software Engineer | Create programs that can not only see visual information, but also interpret it. |
| Data Analyst | A person whose job is to gather and interpret data in order to solve a specific problem. |
| Data Analytics Engineer | Provide clean data sets to end users, modeling data in a way that empowers end users to answer their own questions. Spends more time transforming, testing, deploying and documenting data rather than analyzing data. |
| Data Analytics Lead | Responsible for managing data systems and ensuring information accuracy to support business operations and requirements. Work with system engineers to maintain all the database’s safety and security to prevent unauthorized access and potential leak of confidential data. |
| Data Analytics Manager | Provide direction for a team of data analysts. The build that team, making hiring decisions and deciding where each analyst’s skills will prove most productive for the organization. Work with analytics department, ensuring its accuracy. |
| Data Architect | An IT professional responsible for defining the policies, procedures, models and technologies to be used in collecting, organizing, storing and accessing company information. |
| Data Engineer | Responsible for designing, maintaining, and optimizing data infrastructure for data collection, management, transformation, and access. In charge of creating pipelines that convert raw data into usable formats for data scientists and other data consumers to utilize. |
| Data Engineering Manager | Responsible for managing the many complex components of a modern data stack. Modern data stack is an integrated set of tools that helps facilitate the handling, cleaning, processing, and storing of data. |
| Data Science Consultant | Works with organizations to perform analyses and computations that draw insights from collected data. Work with clients to improve data competencies, analytical skills and business strategies. |
| Data Science Engineer | IT worker whose primary job is to prepare data for analytical or operational uses. Typically responsible for building data pipelines to bring together information from different source systems. |
| Data Science Manager | Responsible for helping the company leverage data, working with team of data scientists and engineers to provide valuable direction and make informed decisions concerning the product, growth, and engagement. |
| Data Scientist | Person employed to analyze and interpret complex digital data, such as usage statistics of a website, especially in order to assist a business. Leads research projects to extract valuable information from big data and is skilled in technology, mathematics, business and communications. |
| Data Specialist | Oversee the development of new databases, monitoring database performance and interpreting raw data and turning it into usable feedback and applications. |
| Director of Data Engineering | Engage in technical discussions with your engineers and managers, working to understand tradeoffs and guide major decisions. |
| Director of Data Science | Establishes, plans, and administers the overall policies and goals of data science function. Provides strategic guidance and overall direction for analytical efforts. |
| ETL Developer | Type of software engineer that manages the Extract, Transform, and Load processes, implementing technical solutions to do so. |
| Finance Data Analyst | Professionals who help financial institutions utilize data to make high-quality business decisions. |
| Financial Data Analyst | Prepares financial reports that serve as summary information for managers. |
| Head of Data | Senior executive responsible for the utilization and governance of data across the organization. |
| Head of Data Science | Consistently identifies and monitors key business risks and realizes the data needs of the business. Manages junior data science teams and oversees all activities ensuring alignment with departmental and business-wide vision and strategies. |
| Head of Machine Learning | Develop and lead a team of brilliant machine learning engineers and data scientists to imaginatively brainstorm new data sources, pipelines, algorithmic approaches and strategies to enhance and expand upon our trend prediction and demand forecasting infrastructure. |
| Lead Data Analyst | Responsible for managing data systems and insuring information accuracy to support business operations and requirements. |
| Lead Data Engineer | Responsible for managing all imaging data, metadata and tools for interacting with such data. |
| Lead Data Scientist | Senior-level professional who manages data projects and leads a team of data scientists within an organization. They understand and analyze data to drive growth and plan and prioritize data projects. |
| Lead Machine Learning Engineer | Designs and develops scalable solutions using AI tools and machine-learning models. Performs research and testing to develop machine learning algorithms and predictive models. |
| Machine Learning Developer | Expert on using data to training models. The models are then used to automate processes like image classification, speech recognition, and market forecasting. |
| Machine Learning Engineer | Professional who specializes in designing and developing machine learning systems. Possess expertise in statistics, programming and data science and their role involves creating self-learning applications. |
| Machine Learning Infrastructure Engineer | Professionals who design, build, deploy and maintain the company’s IT infrastructure. Responsible for making computing systems work efficiently and ensuring the services run smoothly. |
| Machine Learning Manager | Oversee machine learning engineer teams and create devices and systems that can complete an action without human intervention. |
| Machine Learning Scientist | Focus on designing and implementing adaptive algorithms that drive AI systems, working in collaboration with data scientists, data engineers and algorithm specialists. They develop autonomous AI software and conduct tests to ensure that software generates accurate predictions. |
| Marketing Data Analyst | Professional who is responsible for interpreting research about the company’s consumers and their buying habits. |
| ML Engineer | Professional who specializes in designing and developing machine learning systems. Possess expertise in statistics, programming and data science and their role involves creating self-learning applications. |
| NLP Engineer | Engineers that focus on tech’s ability to process and analyze natural language data. Use computer science, information sciences, AI and linguistic skills to create programs that can understand human languages. |
| Principal Data Analyst | This role provides advice on data modeling and analytics approaches based on available data to support development of insights. Responsible for the documentation of the analytics approaches and models employed. |
| Principal Data Engineer | Leads data engineering strategy and delivery across global projects and products. Ensures data engineering function works as one team, promote excellence and technical development and adheres to key principles and processes. |
| Principal Data Scientist | Responsible for leading larger and more complex data science projects than senior data scientists. Expected to be able to handle a high degree of complexity and uncertainty and provide innovated solutions to complex problems. |
| Product Data Analyst | Provides advice on data modeling and analytics approaches based on available data to support development of insights. Responsible for the documentation of the analytics approaches and models employed. |
| Research Scientist | Conducts scientific research or investigation, in order to discover new things. |
| Staff Data Scientist | Provides expertise across analytical/technical aspects of the team including: machine learning, deep learning, data mining techniques, statistical techniques, dimensional reduction, R, GIT and JIRA. |