**DSE5002 Data Project #1**

**Scenario**

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 the 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.

**Deliverables**

Prepare your analysis in an Jupyter notebook . Your final product should be a PowerPoint (or Google Slides) 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. They also want very concise presentations, the total length must be 4 to 6 pages in PowerPoint.

Note: You can add an appendix after the end of the main presentation, if there are details that just don’t fit into the main presentation. The appendix, if used, must be clearly as an appendix. The appendix does not count toward the grade.

Use of an appendix like this in a “real world” presentation allows you to have a highly professional way to answer questions you think your audience may ask by pulling up a table or graphic related to that question.

You will need to turn in the Presentation deck and the code as a Jupyter Notebook in pdf format.

**Data Set**

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) |