Predicting STEM Salaries

Project Goal:

For this project, I would like to utilize my knowledge and skills on regression to predict salaries of STEM workers based on some characteristics related to their employment. For example, company, level, job title, location, and so on.

As a Data Science major student, and working in the tech industry, the project is based on my personal curiosity about the pay of STEM worker from companies such as FAANG. Hopefully, I can know more about some factors that possibly bring higher pay, or get some insights from the data to help my career growth.

Also, I think every STEM specialist likely has the same motivation as me to understand the market pattern, and the market value based on his/her employment situation. I hope this project would help them to make better career decisions.

Proposed Data Set:

In order to proceed the project with Data Science Processes of Get, Explore, Model and Report, I am going to use Kaggle's dataset: Data Science and STEM salaries.

This dataset is publicly available, and containing scraped salary reports from level.fyi, with a few employees' information with following fields:

- Timestamp
- Company
- Level
- Title
- Total Yearly Compensation
- Location
- Years of Experience
- Years at company
- Tag
- Base salary
- Stock grant value
- Bonus
- Gender
- Other details
- Race
- Education