PROJECT PROPOSAL

Problem Statement: What could be a realistic yearly gross salary to be expected for an individual finishing a data science bootcamp within the Kansas City, MO (KCMO) metro area with little experience?

Context: For a new graduate, knowing what your salary will look like can be tough. The tech industry dominates in coastal cities like San Francisco, but what about those roles within the Midwest? Likely the higher cost of living on the coast and skew the data towards higher salaries, but how do you know when a job offer is in line with what is reasonable? This is where this information will become valuable to someone just breaking into the industry in this market and unsure of when to negotiate and when to accept an offer. With better education, better decisions can be made for an individual's future.

Criteria for success: Success will be determined by the ability to clearly tell what the average or expected salary will be for a new graduate, looking for full time work, within the KCMO metro for a handful of data science related roles.

Scope of solution space: We might be limited by being able to find Kansas City specific data, however, we could potentially use algorithms to scale Kansas City pay to US averages to get a more realistic idea of what to expect.

Stakeholders: Project mentor and new data science graduates in the KCMO metro market.

Data sources: We can used data sources on Kaggle that were located and listed below, US employment data available by the US census, and Jupyter notebook.

Additional project information:

- What is the problem you want to solve?
    - What could be a realistic yearly gross salary to be expected for an individual finishing a data science bootcamp within the Kansas City, MO (KCMO) metro area with little experience?
- Who is your client and why do they care about this problem? In other words, what will your client do or decide based on your analysis?

    - Clients could be new data science bootcamp graduates within the Kansas City, MO (KCMO) metro areas to set expectations for what to expect upon graduation and subsequent job searching. This would also provide valuable insight into fair market value to better negotiate salaries and plan financial needs for this specific market as this field can vary widely from location to location and by job title. This will help give a clearer picture of financial outlook and help new graduates avoid under or overvaluing their work, have data to support negotiations, and help guide graduates to job titles that better fit their financial needs.
- What data are you using? How will you acquire the data?
    - There is an interesting data set provided on the Kaggle platform and can be downloaded at : https://www.kaggle.com/datasets/arnabchaki/data-science-salaries-2023
- Briefly outline how you'll solve this problem. Your approach may change later, but this is a good first step to get you thinking about a method and solution.
    - First, I would like to define Kansas City, MO metro area by zip codes within a certain radius of the city or possibly by surrounding counties. As we go into EDA, we can further see which might be a better representation. Next, once we define the area we are wanting to further examine, I would like to look at categories such gross yearly salary, years of experience, and job title. There are other categories, that if the data can be located, I would also like to cross reference with such as gender, work from home status vs. in office, level of degree acquired (bootcamp graduate specifically would be interesting to look at and how this compares to traditional tech degrees), etc.
    - I anticipate that I'm not going to find a specific data set for KCMO but will likely need to scale the gross pay based on location by taking averages of the US pay and how Kansas City measures up to averages within the industry.
- What are your deliverables?
    - Report and slide deck for each section completed along with coding work via Jupyter notebook.