TalentPitch Challenge

Welcome to the TalentPitch Challenge! Good luck, and feel free to reach out to me at if you have any questions.

You will be provided with two tables in `.csv` format. One table (`users.csv`) contains processed user information, while the other (`users_raw.csv`) contains raw user information. Both tables can be joined using the `user_id` and `id` respectively.

The main objective of the challenge is to perform an Exploratory Data Analysis (EDA) on the data:

- **Data Cleaning**: Check for missing values and address them appropriately for subsequent analysis. We recommend using the Missingno library for visualizing missing values. However, be cautious as some fields may appear to contain data but may be empty or invalid.

- **Insight Discovery**: Create tables that showcase interesting insights. Select at least two of the following insights and create one additional insight that you find interesting:

    - Profile completion rate over time.
    - Top 5 roles with the highest `average_feedback` by `gender` .
    - Average `views_to_resume_received` by `city` .
    - Top 5 `dreamt_companies` by `desired_state` and `level_last_study` for users with more than 10 `connections_sent` .
    - Top 3 most common `skills` for users with `average_feedback` greater than 4, grouped by `last_study` .

- **Visualization**: Create at least three of the following plots and develop two additional plots that you find useful or interesting:

    - Stacked bar plot of `forms_to_work` by `level_last_study` .
    - Heatmap of the most common `languages` by `last_study` .
    - Scatter plot of `age` vs. `average_feedback` .
    - Line plot of `profile_completed` over time (by day, by week, etc.).

    - Heatmap of `profile_completed` by `city` and `state`.
    - Stacked bar plot of top 5 `roles` by `gender`.
    - Scatter plot of `views_to_resume_received` vs. `reactions_received`, colored by `profile_completed`.

Submission Requirements

Please submit a repository containing the following files:

- `Dockerfile`: Contains instructions to create a Docker image.
- `docker-compose.yml`: Configuration file for running a multi-container Docker application.
- `requirements.txt`: Lists the necessary libraries to run the project in a Jupyter notebook.

Additionally, include a Jupyter notebook containing all analyses performed and well-documented code.

The completion of all tasks is not mandatory. We will evaluate the quality of your code and your logical thinking in solving the assigned tasks.