Junior Job Opportunities around the United Nations

This is the final report for Junior Job Opportunities around the United Nations visualization project. We create a 3D heatmap with temporal and spatial elements to concretize the employment rates of young migrants in UN member states. Source code is available at Github.

Motivation

To enhance the understanding of entry-level job trends within United Nations member states, this project aims to deliver in-depth insights that will support jobseekers in navigating the job market, aid educators in tailoring their curriculums to better prepare students for these opportunities, and assist policymakers in crafting informed strategies that align with job market realities. These insights are designed to facilitate more effective decision-making processes in adapting to the dynamically changing employment landscape.

Scene

The scene of the visualization is the countries of UN. The countries can be compared and contrasted with each other to arrive at conclusions about their respective positions in the junior job market.

Character

The 'UN Job Seeker'—young professionals entering the job market, facing economic and industrial challenges within UN member states.

Conflict

Job market volatility caused by economic shifts, technological advancements, and geopolitical changes complicates career paths for junior jobseekers.

Approach

- **Data Collection:** Focus on UN member states by gathering employment data from the International Labour Organization (ILO), Department of Economic and Social Affairs, and the World Bank. This includes historical data on job availability, industry sectors, and economic indicators relevant to the labor market.

- **Data Processing:** Employ advanced data analytics and machine learning techniques to clean and analyze the data. This might include removing stop words from natural language data, stemming, normalizing, collecting time series', creating embeddings. Special attention will be paid to socio-economic indicators such as GDP, employment rates across different regions and the amount of junior level positions with respect to population.
- **Visualization Techniques:**
  - **Trend Lines** and **Time Series Analysis** to visualize the fluctuation of job opportunities over time.
  - **Heat Maps** and **Geographical Information Systems (GIS)** to show spatial distribution of job opportunities and economic health within and across UN member states.
  - **Interactive Filters** We include log scale filtering for the bar plots to make the visualization more readable.

Tools

- **Visualization** will be done with pydeck python package. It will be a 3-dimensional heatmap that communicates the user the ratio of junior jobs to all jobs, as well as the GDP of the country. The GDP will be communicated via color from blue to red, where red is equal to the lowest GDP and blue equal to the highest GDP. We will combine pydeck with streamlit that allows us to add interactive widgets.
- **Datasets** are
  - ILO Labour Migration - Source for data on the number of young migrant workers aged 15-24.
  - Per capita GDP at current prices - US dollars - Utilized for GDP data of UN member states.
  - geoBoundaries.org - Provides the GeoJSON data.

Objectives

- **Enhanced Understanding:** Offer a clear depiction of trends.
- **Strategic Planning:** Provide data-driven guidance for the UN jobseeker.

Resolve

To empower UN member states with a robust tool to reduce the employment gap, fostering a more informed and efficient job market in UN member states

1 Ensuring usability

We believe young jobseekers within UN member states will learn which country offers stability and prosperity when doing over-time observation of GDP sizes and junior migrant employment rates, we will know when we see people have used the temporal filtering of the application.

This summarizes the main focus of our application and that we are interested in offering people an overview over the temporal dimension of the state of affairs in the UN member states.

We provide an interactive application that has our dataset, visualization, and filtering options pre-configured for the user. It is a fairly simple use-case that we intend to make as easy as possible for the user to reach.

2 Interactions and time

The user will be able to adjust a slider over the years that are available in the data. The user will see the GDP change as indicated by the color of the bars, as well as the height of the bars change according to the amount of youth migrant workers in the member states. The visualization is a 3D map, and the bars will be on top of the countries. The bars are shaped identically to the countries outlines.

Above the map, you have the filtering section, which will allow you to change the year of the visualization, as well as some basic adjustments like log-scaling the data.

Time commitment depends on the user, by and large. The user may only glance a single year, look at the GDPs and relative youth employment rates, and close the application with useful new insights into the member states. At the other extreme, the user may carefully analyze each country over time, that is, use the temporal slider to observe the changes and fluctuations that the countries have experienced over time. This would take the user up-to a considerable amount of time. We expect the average usage time of the visualization to be 5 minutes.

3 Data discussion

We have observed the data manually and determined the data to be complete up to 2021, which is when the working-age population tracking began by the International Labour Organization. We intend to include in the visualization columns "Value", which is the amount of migrant workers, where columns "Age" has value "Youth, adults: 15-24". This means that by junior migrant workers, we mean 15—24-year-old migrants. The International Labour Organization sources this data from various outlets, depending on the country in question. Additionally, we are using UNdata as a provider of GDP per capita, which we will contrast with the migrant worker population in the visualization. The

currency of the GDP per capita is reported in US dollars. The migrant worker amounts will be reported in thousands of workers.

Temporal Relevance and Approach

Time plays a key role in understanding the trends and fluctuations in junior job opportunities across UN member states. By exploring employment data across various years, users can observe the impact of economic, technological, and geopolitical shifts on job availability for young professionals. This temporal perspective allows users to assess how stable or volatile the job market has been in specific regions, providing insights that would be missed by a single-year snapshot. For instance, significant events like economic recessions or policy changes affecting migration can be observed as clear shifts in the visualized data over time.

Communicating Temporal Aspects

To effectively communicate the role of time, our visualization incorporates a temporal slider, enabling users to explore data year-by-year. By adjusting the slider, users can witness changes in GDP and junior migrant worker counts over time, reflected through dynamic bar heights and color shifts in the 3D map. Each country's bar height reflects the GDP per capita of the country, while the color intensity indicates youth employment levels, making the interplay between economic health and employment easily recognizable.

This time-based interaction supports diverse exploration approaches: users might opt for quick, year-specific comparisons to grasp a country's current job market, or perform a more thorough, multi-year analysis to uncover long-term trends. This flexibility allows users to tailor their experience to their informational needs, making the temporal aspect a valuable layer in understanding junior job market trends within the UN framework.

Visualization and Code

In this section, we will display different snapshots of the visualization in real-world use. After the visualizations, we will detail the code and libraries used to create the visualization.

Visualization

First, let's look at the entire visualization overview. In Figure 2, we can see the elegant design of the front-end of our visualization application. The layout has been chosen to make the visualization as accessible as possible. The interface represents a common 30/70 split, where 30 % of the view is dedicated to controls and information about the visualization, and 70 % is dedicated to the

visualization itself. On the left-hand-size, in the meta-information section, you have controls given to the user. The most important element of our interactive experience is time. Thus, the user can use the slider to select any year from 2000 to 2021, which is the range in which we have data available. The user may observe how the visualization changes each year. In this way, implicit information can be inferred, such as the trajectory of a given UN member's GDP per capita over time. Furthermore, color coding informs the user of the relative state of employment of a given member state. The legend includes all necessary information, such as explanations for different color codings and the source of the data. We also provide a log-scale visualization to even out exponential-scale differences for a cleaner, more even map.

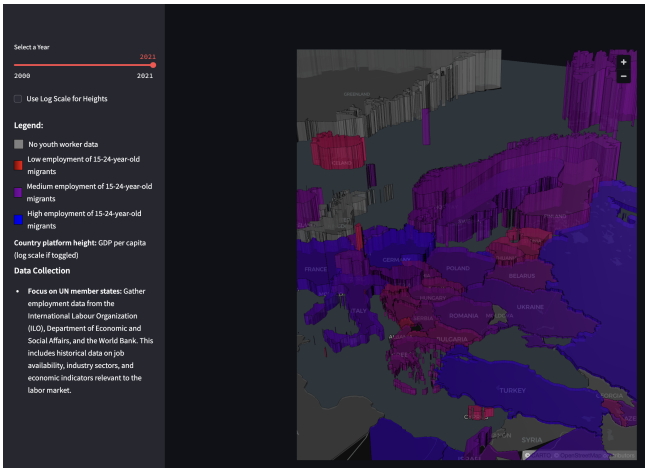

Figure 1: Image of the visualization dashboard. Here, the user can browse a map of the UN countries (other countries included, but not marked with 3D bars or color-codings). In addition, the user can inspect the legend, as well as interactive controls.

Code

In this subsection we will inspect the code and data processing related to our visualization. There is a surprising amount of code required to consolidate multiple datasets and e.g. map countries to each other via arbitrary data vendor specific country codes and/or ISO3166-1 alpha-3 country codes.

We will look at a few highlighted areas of the code. First, the mapping of datasets, and second, the creation of the frontend using a library called streamlit. Moreover, we render the map by using a library called pydeck, on top of which we overlay 3D bar charts based on down scaled GeoJSON files from geoBoundaries.

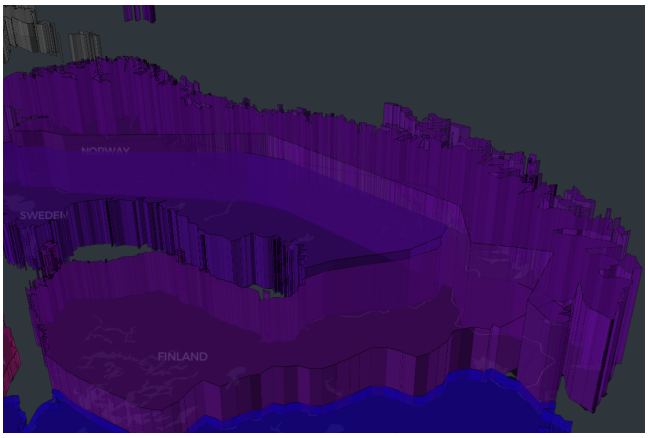

Figure 2: A close up of the best Nordic countries. The user may select any region in which GDP and youth employment is of interest. Here, it is clearly emphasized that Norway has the highest GDP per capita (indicated by bar height), and Sweden has the lowest migrant employment rate (indicated by the most blue color).

Mapping the datasets

In the implemented code, demographic and economic data are processed, focusing on GDP per capita and statistics of migrant youth workers aged 15–24. The pandas library is utilized for data manipulation tasks such as filtering, dropping missing values, and normalizing data. Country code conversions are handled using the pycountry library. Additionally, numpy is used for numerical operations, and JSON for managing GeoJSON file operations. The primary output is an enriched GeoJSON file that integrates detailed economic and demographic metrics, enhancing the dataset for potential analytical or visualization purposes.

Interface Implementation

We use several Python libraries to manipulate data and create the visualizations. The pandas library is used for data processing, while NumPy handles numerical calculations such as the log transform for the bar heights. Country code transformations are managed using pycountry, and JSON processes GeoJSON files for spatial data representation. The visualization layer utilizes pydeck for rendering geospatial data in 3D, and streamlit, along with its components module, to build an interactive web application that allows dynamic user interaction with the data through controls such as sliders and checkboxes.

References

1. Source code - Source code.

2. pandas - A fast, powerful, flexible, and easy to use open-source data analysis and manipulation tool built on top of the Python programming language.
3. NumPy - A fundamental package for scientific computing with Python, including support for a powerful N-dimensional array object.
4. pycountry - A Python library providing the ISO databases for the standards of countries, languages, currencies, and scripts.
5. pydeck - A high-scale WebGL-powered framework for visual exploratory data analysis of large datasets.
6. streamlit - An open-source app framework for Machine Learning and Data Science projects.
7. Streamlit Components - Allows developers to create new Streamlit components or wrap existing React components with Python.
8. ILO Labour Migration Data - Labour migration statistics provided by the International Labour Organization.
9. UNdata GDP per capita - Per capita GDP at current prices - US dollars from United Nations Statistics Division.
10. geoBoundaries - A global database of political administrative boundaries for use in geographic information systems (GIS).