# TEAM 2 – PROGRESS REPORT
## LivabilityX: An Analysis of Trends in Global Livability

Yulin Wei; Danyang Zhang; Xingzhou Ma
weiyulin020626@163.com; {dzhang497;lyndonma}@gatech.edu

### 1. Introduction

Livability significantly impacts residents' quality of life and informs development and policymaking[1]. With challenges from urbanization, global warming, and socioeconomic disparities, understanding and improving livability is crucial[2]. Livability involves factors like economic stability, environmental quality, healthcare, and housing. Our project aims to develop a comprehensive Livability Index to assess countries worldwide, integrating climate and socio-economic data to guide policies and planning. By analyzing both historical and current data, our index tracks changes in livability and assesses the impact of emerging trends.

### 2. Problem Definition

This project aims to create an interactive tool for evaluating global livability using climate and socio-economic data. Current assessments are often limited in scope and lack adaptability to changing conditions. Our proposed tool will identify key factors contributing to livability and their changes over time, using machine learning to predict scores and analyze factors like climate, healthcare, cost of living, and safety. The tool will also simulate potential changes in livability, helping stakeholders improve living conditions and adapt to socio-economic and environmental challenges.

### 3. Literature Survey

Ahmed et al. (2019) critically reviews the concept of "livability," noting its broad application and inherent ambiguity in urban planning. They have concluded by discussing strategies to enhance livability in urban environment[3].

Current livability assessments often rely on separate indices or non-integrated measures, making it difficult to capture overall livability comprehensively. For example, the study by Wang et al. (2015) compared Beijing with other global cities using the Livable Level Integrated Index (LLII) and revealed gaps, particularly in environmental quality, which underscores the need for more holistic assessments that consider multiple factors over time[4]. Kutty et al. (2022) propose a novel machine learning-based framework combining metric-distance analysis and clustering techniques to evaluate resilience and livability in 35 European smart cities[5]. Tan et al. (2016) introduced the Global Livable Cities Index (GLCI), incorporating a broader range of indicators to account for economic competitiveness and diverse populations[6], but it could not still dynamically assess climate change impacts and emerging socio-economic trends.

The study by Saeed et al. (2020) emphasized a more integrated approach by using the Analytical Hierarchical Process (AHP) to develop a composite index that accounts for disparities[7]. Najafi et al. (2024) evaluate livability in District 1 of Tehran using GIS to analyze physical, socio-economic, and accessibility indicators at the block level[8]. However, they were limited to a specific region and did not fully capture global or climate-related factors.

Benita et al. (2021) proposed a Spatial Livability Index for dense urban centers, which identified patterns based on geographical correlations but still did not address the dynamic, evolving nature of livability under climate change[9]. Tran et al. (2021) develop transportation livability-related indicators for a green urban road rating system in Taiwan[10]. However, the study's reliance on local conditions may limit its applicability to other regions, necessitating a more flexible framework to accommodate diverse contexts.

Agbali et al. (2019) explore the implementation of livability strategies through stakeholder perspectives in Manchester, Boston, and San Diego, highlighting the significance of social and technological innovations in addressing urban challenges[11]. Al-Maliki et al. (2024) propose an ICT framework for developing livability in Saudi Arabia, addressing the unique challenges posed by the region's religious and cultural context while identifying potential benefits and obstacles[12]. However, the proposed model may require further refinement to adapt to rapid technological change. Kalenyuk et al. (2024) analyze the impact of digital financial technologies on livability, categorizing innovations like payments, lending, investing, and blockchain while emphasizing the need for significant resources and collaborations to implement smart projects[13], but they caution that these advancements also introduce risks such as cyber threats and data protection challenges. Therefore, current practices are often fragmented, region-specific, or static, lacking a predictive, comprehensive, and adaptable approach that can effectively guide long-term planning and policy-making.

## 4. Methodology

**Dataset Collection**

We've chosen 60 representative countries from all continents worldwide and collected their various aspects of data structured as below:

| Criteria Layer | Factor Layer | Source |
|---|---|---|
| Safety | Crime Rate | WORLD POPULATION REVIEW (WPR for short at below) |
| | Clearance Rate | |
| | Homicide Rates | |
| | Incidence of Sexual Violence | |
| | Robbery Rate | |
| | Police Officers | NationMaster |
| Health Care | Global Health Security Index | |
| | Life expectancy at birth | |
| | Birth rate | |
| | Physicians > Per 1,000 people | |
| | Hospital beds > Per 1,000 people | |
| | Per capita health expenditure | |
| | Nurses and midwives > per 1000 | |
| | Doctors > per 1000 | |
| | Universal Health Coverage Index | WHO |
| Purchasing Power | Average monthly disposable salary > After tax | NationMaster |
| | Rent | WPR |
| | Basic utilities > Garbage, water, heating, electricity for 85 sqm apartment | NationMaster |
| | Groceries | WPR |
| | Restaurants | |
| | Local Purchasing Power | NationMaster |
| Education | High school enrolment rate | |
| | Primary education, teachers per 1000 | |
| | Secondary education, teachers per 1000 | |
| | High School education, teachers per 1000 | |
| | Compulsory education duration | |
| Government | Level of Corruption | TRANSPARENCY |
| | Happiness Index | WorldHappiness |
| Pollution | $CO_2$ Emissions per 1000 | NationMaster |
| | PM10 > Country level > Micrograms per cubic meter | |
| | NOx emissions per populated area | |
| | Total renewable water resources | |
| | Drinking Water Quality | WPR |
| | Garbage Disposal | |
| | Noise & Light Pollution | |
| Climate | Average Temperature | WorldEconomics |
| | Average Dewpoint | UNData |
| | Natural Disaster Risk | WPR |

### Database Building

Upon collecting all datasets, the next step was to clean and merge them into a cohesive and searchable database. The dataset was processed using Python and Pandas, where transformations were performed to standardize the various indicators. The overall formula for calculating the Livability Index is shown below, where each criterion layer term is calculated by a formula consisting of all the factor layer terms from their respective categories, due to space constraints only the overall formula ($I_{main}$) and the Climate Index formula ($I_{cl}$) are shown in this report, the rest of the formulas will be given in the final report and at the code level.

$$I_{main} = w_{sa}I_{sa} + w_{hc}I_{hc} + w_{pp}I_{pp} + w_{ed}I_{ed} + w_{go}I_{go} + w_{po}I_{po} + w_{cl}I_{cl}$$

$$I_{cl} = T_{ave} + 0.555\left(6.1e^{5417.75\left(\frac{1}{273.16} - \frac{1}{DP_{ave}+273.15}\right)} - 10\right) - 0.4ND$$

After cleaning, all fragmented datasets were combined into a single searchable SQLite database which was integrated into a Python Flask backend, enabling efficient access for downstream analytics and visualization purposes. We also precomputed weighted metrics using custom Python scripts to allow for rapid querying. By combining all raw datasets into one structured and normalized database, we ensured the data was ready for machine learning analysis, predictive modeling, and insightful visualizations, ultimately aiding planners and policymakers in making informed decisions.

### Multi-model comparative analysis

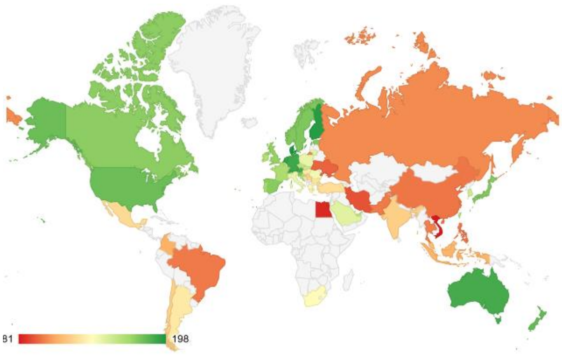
We utilized several machine learning algorithms to predict and evaluate the livability index. The main models employed include Random Forest Regressor, SVM Regressor, and Logistic Regression. Random Forest was particularly useful for handling nonlinear relationships across climate, economy, and social features, while the SVM Regressor effectively captured complex relationships using a radial basis kernel function. Logistic Regression allowed us to discretize the livability index into categories, providing insights into how different factors influence a country's livability ranking.

To ensure objective and reasonable feature weighting, we used the Entropy Weight Method, which calculates the information entropy of each feature to determine its importance. This approach ensures that features with higher variability across countries are given more weight in the livability prediction. The entire data set was normalized before model training to mitigate scale differences and improve prediction accuracy.

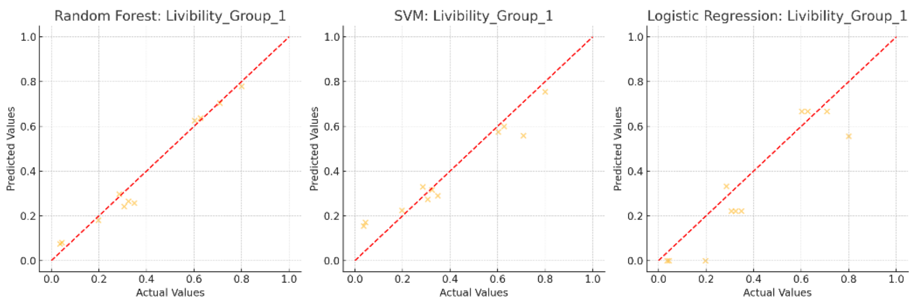
Furthermore, the feature importance analysis and entropy weighting were used to identify the most influential factors affecting livability. This analysis provides valuable data support for making informed policy recommendations to improve livability.

### Visualization

a. Index World Maps: The main index and criteria layer indexes will be presented on world maps. For example, here is the 2017 Quality of Life Index visualization:




b. Regression Prediction Plot: A regression prediction plot is generated that shows how each model (random forest, support vector machine, logistic regression) compares the predicted value of the country's livability index to the actual value. Here below is the group 1 example.




c. (In Construction) Various Aspects of Prediction Maps with Suggestions.

**Innovation List**

a. Multi-model comparative analysis to find the most suitable algorithm for index prediction.
b. The combination of entropy weight method and machine learning.
c. Multi-level regression prediction.
d. Feature importance Analysis and Policy Explanation.

## 5. Experiment and Evaluation

To verify our model's reliability, we will compare its livability index predictions with established sources like the Mercer and EIU indices. This will assess our model's performance and alignment with existing systems. Metrics like Pearson's correlation or Kendall's Tau will quantify the level of agreement, demonstrating its accuracy in reflecting differences in national livability. For key aspects like air quality, population density, and income, we will compare model outputs with real-world data using statistical measures such as MSE and MAPE to identify discrepancies and guide improvements.

We will summarize livability performance among major countries, highlighting strengths and weaknesses. Using radar and heat maps, we will visualize country-specific indicators like air quality and economic standards, while clustering analysis will group countries with similar profiles. This will uncover actionable insights for countries to improve based on similarities with others.

We acknowledge certain limitations. Our weighting system, based on existing research, may introduce subjectivity, which we plan to refine through user surveys and expert feedback. The model's suitability for smaller countries will also be improved through sub-models. We aim to introduce customizable features, allowing users to adjust indicator weights for tailored evaluations.

Finally, a sensitivity analysis will evaluate how variations in indicators affect the total score, helping identify the most influential factors. By systematically adjusting each key variable, we will determine its impact, guiding model optimization and helping users understand the key factors driving livability. This approach ensures high accuracy and practical decision-support value.

## 6. Plan of Activities

| Original Plan | | Yulin | Danyang | Yuchen | Leijun | Revised Plan | Yulin | Danyang | Xingzhou |
|---|---|---|---|---|---|---|---|---|---|
| Activity | **StartDate** | **Duration / hours** | | | | **StartDate** | **Duration / hours** | | |
| Brainstorming & Pick Topic | √ **09-14** | 10 | 10 | 10 | 10 | | | | |
| Proposal | √ **09-16** | 2 | 14 | 2 | 2 | | | | |
| Proposal Slides & Pre | √ **09-20** | 2 | 2 | 14 | 2 | | | | |
| Data Collection & Cleaning | **9-29** | 5 | 34 | 34 | 5 | √ **09-29** | 12 | 45 | 12 |
| Coding | **10-7** | 45 | 15 | 15 | 45 | √ **10-10** | 60 | 30 | 60 |
| Progress Report | **10-21** | 5 | 5 | 5 | 5 | √ **10-21** | 5 | 15 | 5 |
| Final Poster | **11-4** | 8 | 2 | 2 | 8 | **11-1** | 8 | 5 | 8 |
| Final Pre & Report | **11-20** | 25 | 20 | 20 | 25 | **11-13** | 30 | 30 | 30 |
| In Total | | 102 | 102 | 102 | 102 | | 115 | 125 | 115 |

Despite the major changes in the team members due to unexpected events such as the reorganization of the school district during the project, we still tried our best to overcome a series of difficulties and successfully made the planned progress, in which each of the existing team members equally shared the workload more than the original one, and we sincerely thank each of the team members for their unselfish contributions as well as the teacher, Dr. Liu, for her timely and crucial help.

**References:**

[1] Schindler, Mirjam, and Rita Dionisio. "Planning for Urban Livability: Integrating Socio-Spatial Indicators in City-Making." *Journal of Urban Technology* (2024): 1-30.
[2] Ruth, Matthias, and Rachel S. Franklin. "Livability for all? Conceptual limits and practical implications." *Applied geography* 49 (2014): 18-23.
[3] Ahmed, Nora Osama, Amr Mostafa El-Halafawy, and Ahmed Mohamed Amin. "A critical review of urban livability." *European Journal of Sustainable Development* 8.1 (2019): 165-165.
[4] Wang, Jiao, et al. "A comparative study of Beijing and three global cities: A perspective on urban livability." *Frontiers of Earth Science* 5 (2011): 323-329.
[5] Kutty, Adeeb A., et al. "Urban resilience and livability performance of European smart cities: A novel machine learning approach." *Journal of Cleaner Production* 378 (2022): 134203.
[6] Giap, Tan Khee, Woo Wing Thye, and Grace Aw. "A new approach to measuring the liveability of cities: the Global Liveable Cities Index." *World Review of Science, Technology and Sustainable Development* 11.2 (2014): 176-196.
[7] Saeed, Urooj, et al. "An integrated approach for developing an urban livability composite index—A cities' ranking road map to achieve urban sustainability." *Sustainability* 14.14 (2022): 8755.
[8] Najafi, Ehsan, et al. "A GIS-based Evaluation of Urban Livability using Factor Analysis and a Combination of Environmental and Socio-economic Indicators." Journal of Geovisualization and Spatial Analysis 8.2 (2024): 1-20.
[9] Benita, Francisco, Vyacheslav Kalashnikov, and Bige Tunçer. "A spatial livability index for dense urban centers." *Environment and Planning B: Urban Analytics and City Science* 48.7 (2021): 2006-2022.
[10] Tran, Nam Hoai, et al. "Developing transportation livability-related indicators for green urban road rating system in Taiwan." *Sustainability* 13.24 (2021): 14016.
[11] Agbali, Mohammed, et al. "Are smart innovation ecosystems really seeking to meet citizens' needs? Insights from the stakeholders' vision on smart city strategy implementation." *Smart Cities* 2.2 (2019): 307-327.
[12] Al-Maliki, Saeed Q. Al-Khalidi, Ahmed Abdulmutallab Ahmed, and Osman Ali Nasr. "The Establishment of Smart Cities: Existing Challenges and Opportunities–The Case of Saudi Arabia." (2024).
[13] Kalenyuk, Iryna, et al. "FINANCIAL INNOVATIONS IN THE SMART CITY ECOSYSTEM." *Financial & Credit Activity: Problems of Theory & Practice* 1.54 (2024).