Iran J Econ Stud, 2025, 14(2), 399-420

Iranian Journal of Economic Studies

Journal homepage: ijes.shirazu.ac.ir

A Machine-Learning-Based Typology of Sustainable Development Pathways: Evidence from 19 Countries with a Special Focus on Iran

Forough EsmaeilySadrabadi[a*], Samineh Ghasemifar[b]

a. Department of Economics, Ardakan University, Ardakan, Iran.

b. Department of Economics and Accounting, University of Hormozgan, Bandarabbas, Iran.

Highlights

- **A data-driven analysis identifies distinct pathways toward sustainable development and positions Iran among comparable emerging economies.**
- **The findings reveal significant potential for Iran to enhance renewable energy adoption and energy efficiency.**
- **Policy-oriented insights highlight strategic directions for aligning economic growth with environmental sustainability.**

Article History

Received: 22 November 2025
Revised: 13 December 2025
Accepted: 19 December 2025
Published:26 December 2025

JEL Classification

C32
I18
Q20
Q50

Keyword

Load Capacity Factor Energy Intensity
Human Capital Environmental Quality

Abstract

This study develops a novel, fully unsupervised, data-driven typology of national sustainability pathways to identify distinct development trajectories and to position Iran within a global comparative context. The methodological framework combines K-means clustering optimized by the silhouette score, t-SNE visualization for high-dimensional pattern recognition, and hierarchical clustering for robustness validation, providing an objective basis for cross-country sustainability assessment.The analysis relies on five key indicators—GDP per capita, CO_2 emissions per capita, renewable energy share, energy intensity, and the Human Development Index (HDI)—for 19 selected countries over the 2021–2025 period. This multidimensional approach captures core economic, environmental, and social dimensions of sustainable development.The results identify four distinct sustainability pathways. Iran is clustered with fossil fuel–dependent emerging economies, including China, India, Turkey, Malaysia, and Thailand. Within this group, Iran shows the lowest renewable energy share (1.1%) and the highest energy intensity, placing it on the most fossil-fuel-locked trajectory in the sample. In contrast, benchmark countries such as Sweden achieve renewable energy shares above 60%, revealing a substantial sustainability gap.These findings provide policymakers with a clear benchmarking tool and highlight three priorities for Iran: expanding renewable energy investment, reforming fossil-fuel subsidies, and modernizing energy infrastructure to support a sustainable growth path.

1. Introduction

Sustainable development has become a fundamental pillar of contemporary economic progress, particularly for countries facing structural constraints and

* f.esmaeily@ardakan.ac.ir
DOI: 10.22099/ijes.2025.54083.2081

ecological pressures. In the case of Iran, long-standing dependence on hydrocarbon revenues, high energy intensity, limited penetration of renewables, water scarcity, and growing climate-driven risks have weakened economic diversification and constrained improvements in human development outcomes. These structural vulnerabilities are further amplified by sanctions, regional instability, environmental degradation, and seasonal energy shortages, illustrating the urgency of institutional reforms and coordinated policies aligned with the Sustainable Development Goals.

The phenomenon under study—cross-country sustainability performance and Iran's relative position within it—has been widely examined through descriptive assessments and econometric evaluations.

Existing research provides valuable insights into macroeconomic exposure, environmental degradation, and SDG progress; however, it remains limited in its capacity to reveal deeper structural similarities across nations or to identify strategic reference groups that could inform policy benchmarking.

What remains insufficiently explored is a systematic methodology that classifies countries based on multidimensional sustainability indicators using advanced data-driven techniques rather than predefined regional or income-based categorizations. This gap restricts the development of actionable pathways for countries seeking targeted policy learning and evidence-based transition strategies.

To address this research gap, the present study applies an unsupervised machine-learning framework combining K-Means clustering with t-SNE dimensionality reduction to analyze sustainability and socio-economic indicators across 19 countries from 2021 to 2025. The optimal clustering structure is validated using silhouette scores, enabling a robust classification that highlights structural patterns and positions Iran relative to high-performing countries such as Sweden. This approach provides more than a comparative ranking; it offers a diagnostic tool that uncovers the underlying dimensions shaping sustainability trajectories.

The expected contribution of this study lies in generating new empirical insights into how countries align along environmental, social, and economic dimensions of sustainable development, while simultaneously providing a strategic analytical framework for policy design. By integrating machine learning with sustainability assessment, the research advances a predictive, evidence-based tool that can support long-term planning, inform low-carbon transition policies, and enhance the capacity of countries like Iran to pursue resilient and competitive development pathways.

This paper makes two distinct contributions. First, it develops the first fully unsupervised, machine-learning-based typology of sustainable development pathways using K-means clustering (silhouette-optimized), t-SNE visualization, and hierarchical clustering validation on five core indicators (2021–2025). Second, it provides the first rigorous placement of Iran within a global, data-driven sustainability typology — showing that Iran follows the most extreme fossil-fuel-dependent trajectory in the entire sample. These findings offer policymakers a clear benchmarking tool and actionable policy priorities.

2. Literature Review: Economic Sustainable Development

2.1 Historical Evolution of the Sustainable Development Concept

Sustainable development has taken shape over time through a sequence of key theoretical debates and policy interventions. Early discussions trace back to Malthus (1798), who highlighted the tension between population growth and limited natural resources. Concerns about long-term resource scarcity and environmental pressures resurfaced in the twentieth century. The most influential example was the Club of Rome's "Limits to Growth" report (Meadows et al., 1972). Using systems modeling, it warned that unchecked economic expansion could destabilize the global ecological balance.

A decisive turning point occurred with the publication of the Brundtland Report in 1987, which framed sustainable development around intergenerational responsibility—emphasizing present needs without undermining those of the future.This formulation integrated economic growth, social well-being, and environmental protection into a single framework, shaping all subsequent discussions on sustainability.

The Rio Earth Summit in 1992 further institutionalized the concept through Agenda 21, emphasizing global cooperation on environmental governance and sustainable policy-making. Later, the Millennium Development Goals (MDGs) for 2000–2015 introduced measurable global targets focusing on poverty, education, health, and environmental sustainability. Building on these efforts, the Sustainable Development Goals (SDGs) 2015–2030 expanded the agenda to 17 goals and 169 targets, capturing the multidimensional nature of development and acknowledging the interconnectedness of economic, social, and environmental systems.

This progression—from classical resource debates to comprehensive global frameworks—highlights the increasing complexity of sustainability challenges. It also justifies the move toward data-driven, multidimensional analytical methods such as clustering and machine learning, which are better suited to examine the intertwined dynamics of modern sustainable development indicators.

These historical milestones not only shaped global sustainability agendas but also highlight the critical role of institutional frameworks and behavioral factors in guiding countries toward effective implementation of sustainable development policies. Understanding these dynamics is particularly important for countries like Iran, where institutional and socio-behavioral factors influence the adoption of green technologies and policy compliance.

2.2 Theoretical Foundations of Sustainable Development

Several theoretical perspectives help contextualize cross-country differences in sustainability outcomes and guide the interpretation of empirical findings. Among the prominent analytical approaches, the Environmental Kuznets Curve (EKC) suggests a non-linear relationship in which environmental pressure initially rises with income growth and later declines(U-shape). In early development stages, industrial expansion typically increases pollution, but at higher income levels, technological upgrading, structural transformation, and stronger regulations often lead to improved environmental outcomes (Grossman & Krueger, 1995). This perspective helps explain why high-income nations often cluster into groups with lower environmental stress, while resource-dependent or developing economies remain in more carbon-intensive clusters.

A second important perspective is the decoupling hypothesis, which argues that economic growth can become increasingly independent from environmental degradation due to advances in efficiency, renewable energy deployment, and policy reforms (OECD, 2024). Countries that achieve absolute decoupling—reducing emissions while growing economically—typically demonstrate strong investment in technology, environmental governance, and green infrastructure. This framework clarifies why nations with slower technological diffusion or dependence on fossil-fuel-based growth, such as Iran, remain tightly coupled to environmental pressures.

A third foundational framework is the Technological Innovation Systems (TIS) perspective, which emphasizes how innovation capabilities, institutional quality, market formation, and knowledge diffusion shape the speed and direction of sustainability transitions (Bergek et al., 2008). Strong innovation systems support renewable energy adoption, energy efficiency, and industrial upgrading, while weak systems hinder progress toward sustainability. This perspective is particularly relevant for interpreting Iran's position, as gaps in innovation capacity, governance quality, and technology diffusion help explain its placement within global sustainability clusters.

Together, these theoretical foundations provide a coherent lens for interpreting clustering results

Institutional economics highlights that the quality of governance, property rights, and regulatory frameworks strongly shape the ability of nations to achieve sustainable development goals (North, 1990; Acemoglu et al., 2005). Meanwhile, behavioral economics emphasizes how policy design, social norms, and incentives influence environmental decision-making (Thaler & Sunstein, 2008). Integrating these perspectives provides a more comprehensive lens to interpret cross-country clustering results and to understand why some countries, like Iran, lag in sustainability performance.

These theoretical perspectives directly inform the study's methodological choices, guiding the use of clustering and machine learning techniques to identify sustainability pathways and position Iran within global clusters.

2.3 Literature Review

Research on sustainable economic development began conceptually and institutionally in the late 20th century. Brundtland et al. (1987), in the report Our Common Future, laid the theoretical foundations of sustainable development as an integrated framework reconciling economic growth, social equity, and environmental protection. This framework was later institutionalized through the Millennium Development Goals (2000) and subsequently the Sustainable

Development Goals (2015), forming a multidimensional global policy architecture (Sachs, 2015).

Methodologically, early approaches based on linear models and single-variable indicators have gradually given way to multidimensional and computational methods. For example, Ahani & Afshar-Kazemi (2021) used eight World Bank indicators and clustering techniques to assess Iran's sustainable development status from 1996 to 2016, identifying similar countries such as Indonesia, Vietnam, and São Tomé and Príncipe.

Recent advances in computing and data availability have transformed sustainability research. Najati, et al. (2022) applied deep learning models to Iran's annual time-series data (1970–2018) to predict CO_2 emissions, projecting a rise to 850–900 million metric tons by 2023 and highlighting the utility of deep learning for environmental forecasting. Gholami et al. (2023) demonstrated that a transition towards a green economy could significantly enhance Iran's economic sustainability.

In 2024, studies employed more advanced integrative methods. Sayardoost Tabrizi et al. (2024) proposed a hybrid Machine Learning and Network DEA approach to evaluate Iran's petrochemical supply chain efficiency, showing that clustering homogeneous units improved benchmark realism. Moradi et al. (2024) used Random Forest regression and K-means clustering to assess urban livability in Tehran, revealing distinct resident response patterns and key determinants like access to public transport.

The trend continued in 2025 with studies applying clustering at a global scale. Çelik et al. (2025) used K-Means to classify 166 countries based on SDG indicators, validating results with Random Forest and SVM models. Chen et al. (2025) similarly grouped nations using machine learning and cluster analysis according to economic, social, and environmental structures. Khalili et al. (2025) analyzed Iran's industrial energy consumption and CO_2 emissions using an ARDL model and system dynamics, simulating policy impacts until 2051. Kian Poor & Hajian (2025) examined the interplay between Iran's digital economy and human development, identifying governance challenges. Emerging literature further indicates that unsupervised learning reveals hidden relationships among SDGs and regional development patterns (Garcia-Rodriguez et al., 2025).

Within this evolving methodological context, the present study employs K-Means clustering alongside t-SNE to map Iran’s position within global sustainability pathways. This approach aligns with contemporary computational standards and contributes to the expanding field of data-driven sustainability analysis.

3. Theoretical Framework

Building on the theoretical perspectives reviewed in Section 2.2, this section develops an integrated analytical framework that operationalizes sustainable development through SDG-based indicators and unsupervised machine learning techniques.

Sustainable economic development goes beyond the pursuit of economic growth; it represents a balanced approach that integrates economic performance, social equity, and environmental protection (Sachs, 2015). The concept emphasizes policy designs that not only stimulate growth but also mitigate inequality and reduce pressures on natural resources (Brundtland et al., 1987). In this context, the Sustainable Development Goals (SDGs) serve as standardized global indicators that enable cross-country comparison and provide reliable metrics for assessing progress toward sustainability (Vinuesa et al., 2020; Chen et al., 2025).

From an empirical and data-driven perspective, clustering—one of the core methods within unsupervised machine learning—enables the grouping of economies based on similarities across multidimensional indicators (Jain, 2010). Among clustering techniques, the K-Means algorithm is widely employed due to its computational efficiency and scalability, particularly in economic datasets, facilitating the detection of latent structures and complex sustainability patterns (MacQueen, 1967). To enhance visualization of high-dimensional data, dimensionality reduction techniques such as t-SNE allow researchers to project clusters into two- or three-dimensional spaces, making inter-country structural differences visible and interpretable (van der Maaten & Hinton, 2008).

Despite its advantages, the K-means algorithm has inherent limitations that were considered in this study. These include sensitivity to initial centroid placement, which can lead to convergence on local minima; potential scale-dependency, even after normalization, if variables exhibit extreme variances; and the assumption of spherical, equally sized clusters, which may not fully capture irregularly shaped data distributions (Lloyd, 1982; Rousseeuw, 1987). Potential limitations of the clustering method, including sensitivity to initial centroids, scale dependency, and the assumption of spherical clusters, were carefully addressed through repeated runs, Z-score normalization, and silhouette validation to ensure robustness.To mitigate these, the algorithm was run multiple times (10 independent executions) with random initializations, selecting the solution with the lowest within-cluster sum of squares to reduce initialization bias. Additionally, Z-score normalization was applied to address scale differences, and the silhouette score was used not only for optimal cluster selection but also as a validation metric to assess overall cluster quality and separation, ensuring robustness against assumptions of cluster shape.

Recent empirical studies indicate that countries follow distinct development pathways and that data-driven clustering can uncover these trajectories. For instance, Celik et al. (2025) demonstrated that clustering algorithms can classify countries based on SDG performance and socioeconomic indicators while highlighting regional heterogeneity. Similarly, machine learning has been increasingly applied to SDG analyses to help policymakers identify optimal pathways and tailored cluster-based development strategies (Gohr et al., 2025).

In the case of Iran—characterized by heavy reliance on oil revenues and exposure to international sanctions—a data-driven approach is crucial to identify sustainable development pathways under resource constraints and external shocks. Overall, the theoretical foundations of this study rest on three pillars:

Sustainable Development Theory: Focus on harmonizing economic growth, social welfare, and environmental protection (Sachs, 2015; Brundtland et al., 1987).

Unsupervised Machine Learning Approaches: Application of K-Means and t-SNE to detect and visualize development clusters (van der Maaten & Hinton, 2008; MacQueen, 1967).

SDG-Based Data Analytics: Leveraging standardized global indicators to evaluate cross-country sustainability patterns and policy pathways (Celik et al., 2025; Chen et al., 2025; Vinuesa et al., 2020).

This framework underpins the identification of four distinct global sustainability pathways and enables the classification of Iran's development position within these clusters, ultimately informing evidence-based policymaking for a low-carbon, resilient, and efficient economic transition.

4. Methodology

This study adopts a quantitative, descriptive–analytical design to identify sustainable development pathways among selected countries and assess Iran's position within these trajectories. The use of unsupervised machine learning enables pattern recognition directly from data—without reliance on predefined hypotheses—allowing the analysis to reflect empirical conditions rather than theoretical assumptions.

The sample consists of 19 countries purposefully selected to ensure representativeness across diverse economic structures, development levels, geographical regions, and sustainability profiles, while prioritizing data availability to minimize missing values and enhance analytical reliability. Criteria for selection included:

(1) Economic classification: Drawing from World Bank income groups, the sample incorporates high-income economies (e.g., United States, Germany, France, Sweden, United Kingdom, Japan, Netherlands, Italy, Spain, South Korea), upper-middle-income economies (e.g., China, Turkey, Malaysia, Thailand, Saudi Arabia, United Arab Emirates), and lower-middle-income economies (e.g., India, Iran) to capture a spectrum of development stages and resource dependencies.

(2) Geographical diversity: Countries were chosen from multiple regions, including North America (United States), Europe (Germany, France, Netherlands, Italy, Spain, United Kingdom, Sweden), Asia (China, India, Japan, South Korea, Turkey, Malaysia, Thailand), and the Middle East (Iran, Saudi Arabia, United Arab Emirates), to reflect global variations in environmental challenges, policy contexts, and energy systems.

(3) Sustainability and energy profiles: The selection emphasizes a mix of fossil fuel-dependent nations (e.g., Iran, Saudi Arabia), renewable energy leaders (e.g., Sweden), and transitional economies (e.g., China, India) to enable meaningful clustering and benchmarking.

(4) Data availability: Only countries with complete or near-complete time-series data (less than 5% missing values) for all five indicators (GDP per capita, CO_2 emissions per capita, renewable energy share, energy intensity, and HDI) from 2021–2025, sourced from reliable databases (UNDP, World Bank, IEA, Our World in Data), were included to ensure robustness and avoid imputation biases beyond minimal levels. This non-random, purposive sampling strategy aligns with prior international sustainability studies (Çelik et al., 2025; Chen et al., 2025) that use similar cross-sectional designs to explore multidimensional patterns. By focusing on variation rather than exhaustive coverage, the sample mitigates selection bias while providing a balanced framework for identifying structural similarities and differences in sustainable development pathways, with Iran as a focal case for comparative analysis.

The dataset includes key environmental, economic, and energy indicators covering the 2010–2025 period. GDP per capita and the Human Development Index (HDI) were sourced from UNDP and the World Bank. Carbon dioxide emissions per capita and energy intensity were obtained from the International Energy Agency (IEA) and Our World in Data. The share of renewable energy consumption was retrieved from databases published by the IEA and the World Bank (2024). All variables were collected as annual time series, with data for 2024–2025 consisting of projections from established reports and models to ensure forward-looking analysis. Specifically, these projections are derived from: the World Bank's Global Economic Prospects for GDP per capita; UNDP's Human Development Report (2025) for HDI; IEA's World Energy Outlook (2025c) and Global Energy Review (2025a) for CO_2 emissions per capita and energy intensity; and IEA/IRENA databases (Renewable Energy Statistics 2025) for renewable energy share. These sources use econometric modeling, scenario-based forecasting (e.g., under current policies or stated policies scenarios), and trend extrapolation based on historical data up to 2023–2024, incorporating assumptions on economic growth, policy changes, and technological trends. Five-year averages (2021–2025) were computed to smooth short-term fluctuations, mitigate noise associated with pandemic-era shocks, and reflect persistent structural trends.

Table 1. Key Variables Influencing Sustainable Development

Variable	Unit	Description
GDP_pc	USD	Gross domestic product per capita
CO2_pc	Metric tons	Carbon dioxide emissions per capita
Renewable_share	Percent (%)	Share of renewable energy consumption in total energy use
Energy_Intensity	Energy unit per GDP	Energy consumption per unit of economic output
HDI	0–1 scale	Human Development Index

***Source**: Author's research*

Data preprocessing was a crucial step to ensure reliability, internal consistency, and model readiness of the dataset. All variables were compiled as annual time series and then transformed into five-year averages to smooth short-term fluctuations and reduce the influence of temporary shocks. Missing observations—affecting less than 5% of the dataset—were imputed using the mean of adjacent years to preserve temporal continuity while avoiding artificial bias.

Given that the variables operate on heterogeneous scales (for example, GDP per capita measured in thousands of USD compared to HDI measured on a bounded 0–1 scale), standardization was necessary before applying the clustering algorithm. The Z-score normalization method was employed to ensure comparability across indicators and to prevent variables with large numerical ranges from dominating the distance metrics in K-Means classification.

The Z-score transformation is defined as follows:

$$\boldsymbol{Z_{ij}} = \frac{\boldsymbol{X_{ij} - \mu_j}}{\boldsymbol{\sigma_j}} \tag{1}$$

Expanding on the standardization procedure, let X_{ij} denote the value of variable j for country i, μ_j,represent the sample mean of variable j , and $\sigma_j$denote its standard deviation. This transformation places all variables on a uniform scale and eliminates distortions caused by differing measurement units—such as USD versus metric tons of CO_2 emissions—ensuring that no variable disproportionately influences the clustering outcome.

To identify the most appropriate number of clusters (k), the Silhouette Score criterion was employed. This metric evaluates clustering quality by comparing intra-cluster cohesion with inter-cluster separation, making it particularly useful when no prior theoretical assumption exists regarding cluster structure. The Silhouette Score for each observation i is defined as:

$$\boldsymbol{s(i)} = \frac{\boldsymbol{b(i) - a(i)}}{\boldsymbol{max\,(a(i), b(i))}} \tag{2}$$

Continuing the clustering procedure, a(i) represents the average intra-cluster distance for observation i, capturing cohesion, while b(i) denotes the average distance from iii to the nearest alternative cluster, capturing separation. Values close to **+1** indicate strong cluster separation and internal consistency, values near **0** suggest weak boundaries, and negative values imply cluster overlap and potential misclassification. By computing the Silhouette Score for values of k ranging from 2 to 10, the maximum score—approximately 0.40—was obtained at k=4, which was found to best represent the underlying cluster structure.

The final clustering was performed using the K-Means algorithm, one of the most widely used and computationally efficient unsupervised learning techniques for partitioning multivariate data into distinct groups based on similarity patterns. The algorithm repeatedly reallocates observations across groups with the aim of reducing overall dispersion within clusters, measured by the sum of squared distances. Accordingly, the optimization problem seeks to:

$$arg\,min \sum_{j=1}^{k} \sum_{i \in C_j} \parallel X_i - \mu_j \parallel^2 \quad (3)$$

In this formulation, k denotes the number of clusters, C_j represents the set of countries assigned to cluster j, X_i is the feature vector for country i, and μ_j refers to the cluster centroid, computed as the mean vector of observations within each cluster. The algorithm begins by randomly initializing centroid positions, followed by iterative assignment of countries to the nearest centroid based on Euclidean distance. After each assignment step, centroids are recalculated until convergence is achieved.

In this study, convergence was defined as either reaching a maximum of 300 iterations or achieving a centroid shift below a tolerance threshold of 0.0001, ensuring both computational efficiency and numerical stability. To mitigate the risk of converging to suboptimal local minima—given that K-Means is sensitive to initial centroid placement—the algorithm was executed multiple times (10 independent runs), and the optimal solution was selected based on the lowest within-cluster sum of squares.

To visualize the complex relationships across countries and their corresponding clusters, the t-Distributed Stochastic Neighbor Embedding (t-SNE) algorithm was employed as a nonlinear dimensionality-reduction method that projects high-dimensional data into a two- or three-dimensional feature space. The technique preserves local neighborhood structure by modeling pairwise similarities in the high-dimensional space and reconstructing them in the reduced space using probability distributions based on the t-Student kernel. This enables meaningful visual interpretation of latent patterns and spatial separation across clusters.

t-SNE was implemented with a perplexity value of 30, balancing local and global structure in the data, and a learning rate of 200, ensuring stable optimization of the embedded manifold. The visualization served as a complementary diagnostic to validate clustering results and assess structural consistency among countries.

Following dimensionality reduction, the mean values of all sustainability indicators were computed at the cluster level to characterize each developmental pathway. Iran's positioning within the identified clusters was subsequently evaluated and compared both to countries with similar development trajectories (such as China and India) and to advanced, sustainability-leading economies (such as Sweden). This comparative interpretation provides a nuanced understanding of Iran's relative progress and developmental constraints.

Overall, the methodological framework is data-driven, reproducible, and well-suited to capturing nonlinear dynamics underlying sustainable development outcomes. By moving beyond conventional linear econometric models, this approach uncovers emergent global patterns that more accurately reflect the multi-dimensional nature of sustainability transitions.

5. Findings and Discussion

Using an unsupervised machine learning approach, the present study clustered 19 selected countries based on key economic, environmental, and energy indicators over the period 2021–2025. The variables included GDP per capita, CO_2 emissions per capita, renewable energy share, energy intensity, and the Human Development Index (HDI). After standardizing all variables, the K-means algorithm was applied with the optimal number of clusters determined to be four, based on the Silhouette Score criterion.

The clustering results reveal four distinct groups of countries, each reflecting a unique developmental trajectory that integrates economic performance and environmental sustainability. Determining the preferred cluster count through the Silhouette Score ensures methodological rigor by replacing subjective assumptions with a quantitative selection process. As shown in Figure 1, the Silhouette Score reaches its maximum value at k = 4, confirming high inter-cluster separation and internal cohesion at this configuration.

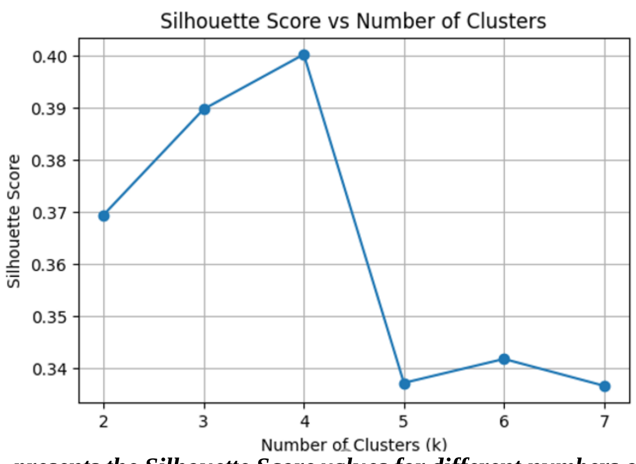

Figure 1. presents the Silhouette Score values for different numbers of clusters (k), illustrating why four clusters provide the most robust and interpretable grouping structure for the dataset.
Source: *Author's calculations*

The final clustering results, obtained through the K-means algorithm and visualized using the t-SNE technique, are presented in Figure 2. This visualization clearly identifies four distinct groups of countries in a two-dimensional space, each representing a unique economic-environmental development pathway. The relative positions of the countries in this embedded space reveal both fundamental similarities within clusters and substantial differences across clusters, highlighting diverse trajectories toward sustainable development.

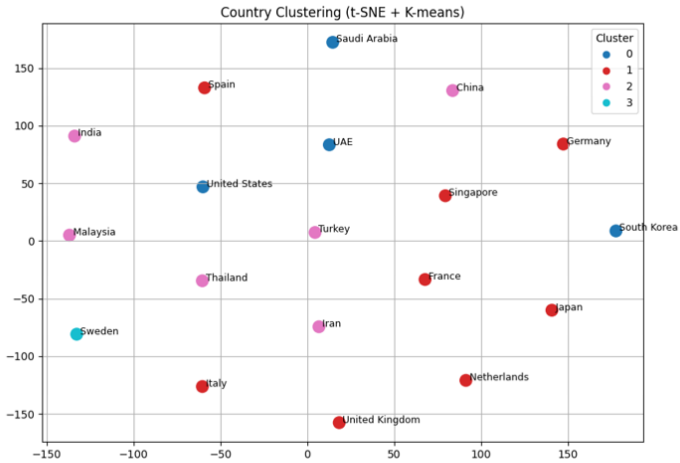

Figure 2. Country clustering using t-SNE and K-means.
Source: Author's calculations*.*

To conduct a more detailed analysis of the characteristics of each cluster, the mean values of key indicators were calculated for all groups. These results are summarized in Table 2, which delineates four distinct development pathways based on a combination of economic growth, environmental performance, and human development.

Table 1. Average key indicators for each cluster

Cluster	GDP per capita (USD)	CO_2 per capita (ton)	Renewable energy share (%)	Energy intensity	Human Development Index (HDI)
0	45,716	18.99	6.59	7,531	0.898
1	47,789	6.73	14.14	3,338	0.942
2	7,293	6.39	15.15	1,321	0.743
3	57,133	3.64	61.7	4,850	0.955

Source*: Author's research*

To further interpret the structural differences across the identified pathways, the average values of key sustainability indicators were computed for each cluster. These aggregated measures reveal distinct developmental profiles: Cluster 0 represents high-income economies with elevated carbon emissions and strong human development; Cluster 1 includes upper-middle-income countries with moderate environmental footprints; Cluster 2 consists of lower-income economies with limited renewable capacity; and Cluster 3 reflects advanced economies with both high income and high renewable energy penetration.

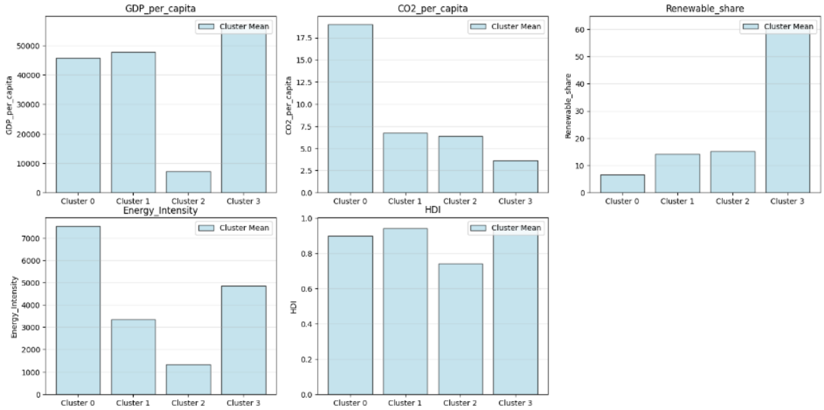

Figure 3. Average sustainability and economic indicators across clusters (2021–2025)
Source: *Author's calculations.*

Figure 3 visualizes the cluster-level averages across GDP per capita, CO_2 emissions, energy intensity, renewable energy share, and HDI. The results illustrate sharp contrasts in energy efficiency and environmental performance. Notably, Cluster 3 shows high income levels combined with low carbon intensity, suggesting a decoupling pattern, while Cluster 0 maintains high emissions despite strong economic output.

Cluster 0 consists of high-income, high-carbon countries such as Saudi Arabia, the United Arab Emirates, the United States, and South Korea. Despite their relatively high GDP per capita and Human Development Index (HDI), these countries face significant environmental challenges. They exhibit very high average CO_2 emissions per capita (around 17 tons) and a very low share of renewable energy (approximately 6%). Energy intensity in this group is also extremely high (average above 7,300 units), indicating a development pathway that remains heavily reliant on fossil fuels and is currently not aligned with sustainable development goals .

Cluster 1 includes developed European countries such as Germany, France, the Netherlands, Italy, Spain, the United Kingdom, and Japan. This group achieves a high GDP per capita (around USD 48,000) and a very high HDI (average 0.94) while successfully reducing CO_2 emissions per capita (around 6 tons) and increasing the share of renewable energy (over 15%). Energy intensity is relatively low in this cluster (average around 3,400 units), demonstrating that economic growth can coexist with environmental preservation (Sachs, 2015).

Cluster 2 comprises developing countries including China, India, Turkey, Malaysia, Thailand, and Iran. These countries have a relatively low share of renewable energy (around 10%) and moderate CO_2 emissions per capita (average 8 tons). GDP per capita is lower compared to other clusters (around USD 10,000), and HDI is at the upper end of the medium-development range (average 0.75). Energy intensity is variable but generally higher than in Cluster 1. Iran belongs to this cluster, reflecting its position as a developing country heavily dependent on fossil fuels.

Iran exhibits the lowest renewable energy share within Cluster 2, at approximately 1.1%, which is well below the cluster average of about 15% and considerably lower than peer countries such as India (35%) and China (16%), as well as Turkey (13%), Thailand (12%), and Malaysia (8%). This pattern reflects a combination of structural, policy-related, and external factors. Historically, fossil-fuel subsidies have shaped price signals and investment incentives, while abundant domestic oil and gas resources with relatively low extraction costs have limited economic motivation for renewable energy adoption (IEA, 2025a; World Bank, 2024). Moreover, international sanctions have constrained access to renewable technologies, project financing, and foreign direct investment in clean energy (IEA, 2025d).

A closer comparison within Cluster 2 highlights key differences. Countries like China and India have benefited from sustained state-led renewable energy programs, technology transfer, and domestic manufacturing development over the past 15 years, enabling higher adoption rates. In contrast, Iran's support frameworks for renewables—such as feed-in tariffs and renewable portfolio standards—have been introduced relatively recently, which helps explain the slower uptake.

Examining additional indicators, Iran's energy intensity (2,950 units) is higher than some cluster peers, including Malaysia (1,735 units) and Thailand (1,430 units), indicating lower efficiency in energy consumption. Per capita CO_2 emissions (8.1 tons) are slightly above the Cluster 2 average (8 tons), highlighting environmental pressures, while the Human Development Index (HDI) of 0.80 places Iran at the upper edge of the medium-development group. This suggests the presence of human capital and social capacity that could support a faster transition toward renewable energy under more favorable policy and investment conditions.

Overall, this within-cluster analysis underscores the structural and policy-related barriers that have limited Iran's renewable energy development, while also pointing to potential opportunities for accelerating the transition through targeted reforms and strategic investments.

Cluster 3 consists solely of Sweden, which demonstrates a unique sustainable development pathway. With 62.9% of energy derived from renewables, very low per capita CO_2 emissions (3.4 tons), and a very high HDI (0.96), Sweden illustrates that achieving economic growth alongside substantial reductions in greenhouse gas emissions and improvements in human development is feasible (Gohr et al., 2025).

To verify the stability of the clustering results, we repeated the K-means algorithm with $k = 3$ and $k = 5$. Iran's cluster membership remained unchanged, confirming the stability of the identified cluster assignments. To validate the stability of the K-means results, we performed agglomerative hierarchical clustering using Ward's linkage on the same standardized dataset. The dendrogram (Figure ۴) confirms a highly similar grouping structure. When cutting the dendrogram at four clusters — consistent with the optimal K-means solution — Iran remains in the same cluster as China, India, Turkey, Malaysia, and Thailand (fossil-fuel-dependent emerging economies).

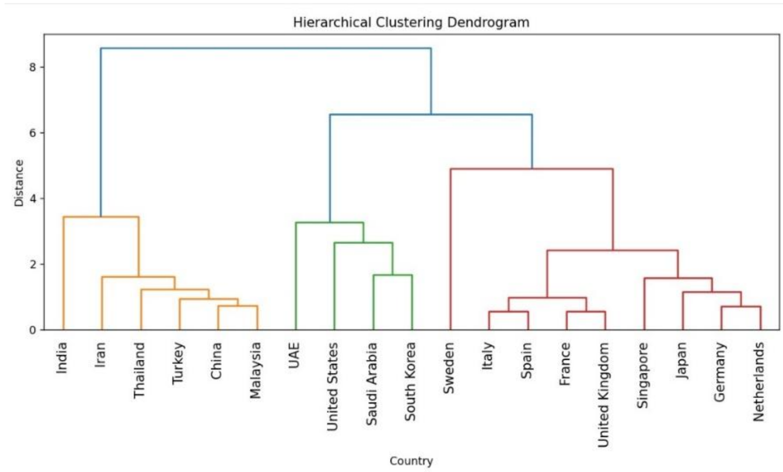

Figure 4. Hierarchical clustering dendrogram (Ward's linkage, Euclidean distance) of the 19 countries based on five standardized sustainability indicators (2021–2025 average)

Source: *Author's calculations.*

This cluster is characterized by low renewable energy share (average ≈10%), moderate-to-high energy intensity, and medium-high HDI. The consistency across both parametric (K-means) and non-parametric (hierarchical) methods strengthens confidence in the robustness of the identified development pathways and Iran's position therein.

These findings indicate that Iran is currently on a development trajectory in which economic growth is accompanied by increased fossil fuel consumption and carbon emissions. This pathway poses environmental, energy security, and economic competitiveness challenges. Transitioning toward sustainable development will require targeted policies to enhance the share of sustainable energy, improve energy efficiency, and reduce carbon intensity. Such measures will not only support environmental protection but also enhance Iran's economic competitiveness and guide the country toward a more resilient and sustainable development model. These results provide a data-driven foundation for Iranian policymakers to design strategies for a low-carbon economy and effective urban energy management (Sadabadi et al., 2025; Heidary et al., 2025).

6. Discussion

The four clusters identified in this study provide empirical validation and refinement of several core theoretical frameworks introduced in Section 2.

Cluster 0 (Saudi Arabia, UAE, United States, South Korea) represents a high-income, high-carbon pathway that directly confirms the persistence of fossil-fuel lock-in even at very high income levels, challenging the classic Environmental Kuznets Curve (EKC) expectation of an automatic turning point after a certain income threshold (Grossman & Krueger, 1995). This finding aligns perfectly with

Sachs' (2015) warning that natural resource abundance and rent-seeking behaviour can perpetuate carbon-intensive growth regimes indefinitely, and extends the analysis of petro-states to non-Middle Eastern cases such as South Korea and the United States.

Cluster 1 (Western Europe + Japan) and the singleton Cluster 3 (Sweden) together provide strong empirical support for absolute decoupling and the downward-sloping portion of the EKC. These economies have achieved high or very high GDP per capita while steadily reducing per-capita CO_2 emissions and dramatically increasing renewable energy shares — exactly the pattern predicted by the decoupling hypothesis (OECD, 2024) and illustrated by Sachs (2015) in The Age of Sustainable Development. Sweden's unique position as a single-country cluster further validates Gohr et al.'s (2025) argument that strong innovation systems, absence of fossil-fuel subsidies, and long-term policy coherence can produce near-complete alignment between economic prosperity and environmental sustainability.

Cluster 2, which includes Iran alongside China, India, Turkey, Malaysia, and Thailand, provides a nuanced perspective on the standard Environmental Kuznets Curve (EKC) narrative. While countries in this group continue to exhibit a strong linkage between economic activity and fossil fuel use, Iran displays a comparatively lower share of renewable energy (approximately 1.1%). This pattern suggests that factors such as long-standing energy subsidies, abundant domestic hydrocarbon resources, and prolonged international sanctions may contribute to a slower transition toward cleaner energy sources, potentially delaying the EKC turning point (Khalili et al., 2025; World Bank, 2024). In contrast, the more rapid expansion of renewable energy capacity in China and India underscores the importance of sustained industrial policy support and access to global technology markets in shaping energy transition pathways.

At the domestic level, Iran's position is broadly consistent with earlier Iranian scholarship (Ahani & Afshar-Kazemi, 2021), which highlights the roles of energy subsidies, industrial structure, and governance arrangements in shaping sustainability outcomes. The clustering results place these country-specific factors within a broader comparative framework, indicating that Iran follows a distinct developmental pattern within the group of emerging economies, rather than simply reflecting a uniform delay relative to higher-performing countries.

Geopolitical constraints further amplify these structural barriers. Prolonged sanctions have restricted access to renewable technology, project finance, and foreign investment — mechanisms that China and India used aggressively during the 2010s. This finding extends the international political economy literature showing how external shocks can lock countries into undesirable development pathways (Bradshaw, 2014; IEA, 2025d).

Taken together, the clustering results move beyond description to actively confirm the EKC and decoupling theories for advanced economies, refine them for resource-rich and sanctioned middle-income countries, and underscore the pivotal role of policy choice, institutional quality, and geopolitical context in determining sustainability transitions — precisely the synthesis called for by Sachs (2015) and the OECD decoupling framework.

7. Concluding Remarks and Policy Recommendations

This study applied unsupervised machine learning to examine the sustainability trajectories of selected countries and to identify Iran's relative position within these global patterns. By integrating K-means clustering with t-SNE dimensionality reduction, the research provided a multidimensional perspective on economic performance, environmental outcomes, and human development. The results show that Iran currently follows a development pathway characterized by high energy intensity, limited renewable energy adoption, and medium levels of human development. This combination reflects a long-standing dependence on fossil fuels and a set of institutional, technological, and policy barriers that restrict progress toward a low-carbon, resilient, and competitive development model.

The analysis highlights the structural nature of the sustainability challenges facing the country. Factors such as coordination across institutions, the pace of digitalization in the energy sector, and the scope of renewable energy strategies appear to influence the speed of progress. Consistent with the international literature, the findings suggest that effective energy transitions are typically associated with sustained policy alignment, technological capability, and integration into global innovation networks. Geopolitical constraints may also shape the extent to which countries can access new technologies, diversify energy sources, and modernize infrastructure. Taken together, these dynamics help explain differences in sustainability performance between Iran and higher-performing economies over time.

Like any empirical investigation, this study is subject to limitations. The scope of countries included in the sample and the reliance on five-year averages may conceal short-term variations or reforms that occurred during the period. The clustering approach, while powerful, cannot fully capture institutional dimensions such as governance quality, innovation capacity, or policy coherence, which are increasingly central to sustainable development. Future research that incorporates governance indicators, broader SDG dimensions, or dynamic clustering methods could deepen understanding of the complex mechanisms driving sustainability differences across countries.

Despite these limitations, the results offer relevant insights for development strategy. Moving toward a more sustainable trajectory may benefit from a longer-term focus on expanding renewable energy capacity, upgrading energy infrastructure, and enhancing institutional coordination. Gradual adjustments to energy pricing and support mechanisms, alongside greater emphasis on clean technologies, could support innovation and economic resilience. In addition, continued progress in energy-sector digitalization, smart grid deployment, and research and development in energy technologies may help reduce existing gaps with higher-performing economies. Developing an integrated framework for monitoring sustainability outcomes could further assist policymakers in formulating informed and adaptive policy responses to economic, environmental, and climate-related challenges.

In moving toward sustainability, Iran stands at a critical juncture. The clustering results show that substantial opportunities exist to shift from a fossil-fuel-dependent trajectory toward a more resilient, technologically dynamic, and environmentally responsible development model. With coherent policies, targeted investments, and strengthened institutional capacity, Iran can gradually converge with the sustainability pathways of high-performing economies while improving long-term economic security and societal well-being.

Author Contributions

Conceptualization, methodology, and formal analysis were conducted by both authors. The first author prepared the original draft, while data collection, resources, and manuscript review and editing were carried out jointly. Both authors have read and approved the final manuscript.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data used in this study were obtained from publicly available international sources, including the World Bank, the United Nations Development Programme (UNDP), the International Energy Agency (IEA), Our World in Data, and IRENA. All data are accessible through the official websites of these institutions.

Acknowledgements

I acknowledge that artificial intelligence tools (including ChatGPT, Grok, and Qwen) were used solely for English language editing and stylistic refinement in this article.

References

Acemoglu, D., Johnson, S., & Robinson, J. A. (2005). *Institutions as a fundamental cause of long-run growth*. In P. Aghion & S. Durlauf (Eds.), Handbook of Economic Growth (Vol. 1A, pp. 385–472). Elsevier.

Ahani, A., & Afshar-Kazemi, M. (2021). Cluster analysis of Iran’s position in global sustainability indices. *Iranian Journal of Development Studies*, 9(3), 45–67. [In Persian].

Bradshaw, M. (2014). *The global political economy of energy*. Oxford University Press.

Bergek, A., Jacobsson, S., Carlsson, B., Lindmark, S., & Rickne, A. (2008). Analyzing the functional dynamics of technological innovation systems: A scheme of analysis. *Research Policy*, 37(3), 407–429.

Brundtland, G., et al. (1987). Our common future. *United Nations World Commission on Environment and Development*.

Çelik, S., Öztürk, Ö. F., Akkuçuk, U., & Şaşmaz, M. Ü. (2025). Global sustainability performance and regional disparities: A machine learning approach based on the 2025 SDG Index. *Sustainability*, 17(16), Article 7411. https://doi.org/10.3390/su17167411

Chen, Y., Li, Z., & Wang, X. (2025). Integrating clustering algorithms and machine learning to optimize regional snapshot municipal solid waste management for achieving sustainable development goals. *Environmental Advances*, 19, Article 100597. https://doi.org/10.1016/j.envadv.2024.100597

García-Rodríguez, A., Núñez, M., Robles Pérez, M., Govezensky, T., Barrio, R. A., Gershenson, C., Kaski, K. K., & Tagüeña, J. (2025). Sustainable visions: Unsupervised machine learning insights on global development goals. *PLOS ONE*, 20(3), Article e0317412. https://doi.org/10.1371/journal.pone.0317412

Gholamia, M., Sadeghi, Z., Jalaee Esfandabadi, S. A. M., & Nejati, M. (2023). Economic and environmental effects of energy transition in Iran by 2050: A dynamic multiregional computable general equilibrium model. *Iranian Journal of Economic Studies*, 12(2), 551–584. https://doi.org/10.22099/ijes.2025.52176.1992.

Grossman, G. M., & Krueger, A. B. (1995). Economic growth and the environment. *Quarterly Journal of Economics*, 110(2), 353–377. https://doi.org/10.2307/2118443

Gohr, C., Rodríguez, G., Belomestnykh, S., Berg-Moelleken, D., Chauhan, N., Engler, J.-O., Heydebreck, L. V., Hintz, M. J., Kretschmer, M., Krügermeier, C., Meinberg, J., Rau, A.-L., Schwenck, C., Aoulkadi, I., Poll, S., Frank, E., Creutzig, F., Lemke, O., Maushart, M., … von Wehrden, H. (2025). Artificial intelligence in sustainable development research. *Nature Sustainability*, 8, 970–978.

Kian Poor, S , & Hajian, M. (2025). Development of digital economy, human development solution in Iran. *International Journal of New Political Economy*, 6(1),47-82.

Heidary, B., Kiani, M. A., & Golzar, F. (2025). Toward sustainable development: Energy transition scenarios for oil-dependent countries, with Iran as a case study. *Energies*, 18(10), 2651. https://doi.org/10.3390/en18102651

International Energy Agency. (2025a). Global Energy Review 2025. IEA. https://www.iea.org/reports/global-energy-review-2025

International Energy Agency. (2025b). Middle East energy outlook. IEA.

International Energy Agency. (2025c). World Energy Outlook 2025. IEA. https://www.iea.org/reports/world-energy-outlook-2025

International Energy Agency. (2025d). Change management in Iran's energy policies: Analyzing the impacts of sanctions and climate change. Energy

Strategy Reviews, 45, Article 101115. https://doi.org/10.1016/j.esr.2025.101115

International Renewable Energy Agency (IRENA). (2025). Global landscape of energy transition finance 2025. IRENA.

International Renewable Energy Agency (IRENA). (2025). Renewable energy statistics 2025. IRENA. https://www.irena.org/Publications/2025/Jul/Renewable-energy-statistics-2025.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651–666.

Khalili, A., Bahmani, M., & Nejati, M. (2025). A dynamic analysis of the environmental impacts and energy consumption of the industrial sector on Iran's economic growth. *Iranian Journal of Economic Studies*, 13(2), 343–368.

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), 129–137.

MacQueen, J. (1967). *Some methods for classification and analysis of multivariate observations*. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability , 1, 281–297.

Malthus, T. R. (1798). *An Essay on the Principle of Population, as It Affects the Future Improvement of Society*. London: J. Johnson.

Meadows, D. H., Meadows, D. L., Randers, J., & Behrens III, W. W. (1972). *The Limits to Growth*. New York: Universe Books.

Moradi, Z., Ziyari, K., & Rajaei, S. A. (2024). Analysis of livability in deteriorated urban fabrics of Tehran's District 7 using machine learning algorithms. *Sustainable City*, 8(2), 1–23. https://doi.org/10.22034/jsc.2024.427202.1748

Najati, M., Sadeghi, Z., & Arabpour, M. A. (2022). Prediction of CO_2 emissions in Iran using key economic indicators and deep learning models. *Environment and Interdisciplinary Sustainable Development*, 7(75), 1–17. https://doi.org/10.22034/envj.2022.149127

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press.

OECD. (2024). Energy Transition Indicators: Policy frameworks and international benchmarks. *OECD Publishing*.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65.

Sadabadi, A. A., Shahi, N., & Borhani, F. (2025). Change management in Iran's energy policies: Analyzing the transition process from fossil fuels to renewable energies. *Energy Conversion and Management: X*, 28, Article 101269. https://doi.org/10.1016/j.ecmx.2025.101269

Sachs, J. D. (2015). The age of sustainable development. Columbia University Press.

Sayardoost Tabrizi, S., Yakideh, K., Moradi, M., & Ebrahimpour, M. (2024). Clustering with machine learning and using NDEA in development planning:

A case study in the petrochemical two-stage SSC. *Research in English Journal*, 5(2), 192–212. https://doi.org/10.22105/riej.2024.455576.1443

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness.* Yale University Press.

United Nations. (1992). Agenda 21: Programme of action for sustainable development. https://sustainabledevelopment.un.org/outcomedocuments/agenda21

United Nations. (2000). United Nations Millennium Declaration. https://www.un.org/millennium/declaration/ares552e.htm

United Nations Development Programme (UNDP). (2025). Human Development Report 2025. UNDP. https://hdr.undp.org/content/human-development-report-2025

Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research,* 9, 2579–2605.

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., … Nerini, F. F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications*, 11(1), Article 233. https://doi.org/10.1038/s41467-019-14108-y

World Bank. (2024). World Development Indicators. https://data.worldbank.org/indicator

World Bank. (2025). Global Economic Prospects, June 2025. World Bank Group. https://www.worldbank.org/en/publication/global-economic-prospects

World Economic Forum. (2025). Fostering effective energy transition 2025. World Economic Forum.