Global Happiness: A Comprehensive Analysis of Social, Economic, and Environmental Factors

Inês Lameira[1] and Romeu Fernandes[1]

[1]Informatics Department, University of Minho, 4710-057 Braga, Portugal.

Contributing authors: pg40082@uminho.pt; pg45861@uminho.pt;

Keywords: data analysis, economic factors, environmental factors, social factors, well-being

1 Introduction

The concept of happiness changed drastically over the past decades, moving from a physiological notion to a measurable and ultimately attractive indicator for social advancement and human development. Such a shift challenges the long-cherished belief that economic growth is the only driving force for societal improvement [1]. The World Happiness Report, an annual collaboration among Gallup, the Oxford Wellbeing Research Centre, the UN Sustainable Development Solutions Network, and WHR's Editorial Board, has shown that global happiness is a compound concept, determined by such multi-faceted dimensions like social relationships, economic security, environmental conditions, and psychological well-being. The evolution of happiness research has gone hand in hand with the increasing awareness of the limitations of conventional economic measures such as GDP. While such indicators do a good job of measuring material wealth, they do not account for influential components of human experience and quality of life, which play a major part in determining well-being [2]. This has prompted the inclusion of broader indicators in measuring societal progress, including social support, perceived freedom to make life choices, trust in institutions, and environmental sustainability, in addition to standard economic measures. Studies further show that higher levels of income produce diminishing returns to happiness, above the subsistence level needed to meet basic wealth needs—a phenomenon referred to as the Easterlin Paradox [3]. This proves the necessity for the making of decisions seeking to bestow non-economic factors that could help in fostering overall wellbeing. Further, cultural influences determine the way in which an individual perceives and expresses the phenomenon of happiness. Whereas some cultures may emphasize social harmony, others place a premium on personal accomplishment [4]. Such differences render it essential for a more calibrated view of happiness and a happier life that accommodates cultural diversity with universal trends.

Since the United Nations declared happiness the basic human aspiration and established the World Happiness Report in 2012, research on global happiness has gathered momentum. This marked forays into research and policy efforts toward figuring out the factors influencing well-being at both the individual and societal levels [1]. Happiness research encompasses different fields that include economics, psychology, environmental science, sociology, and neuroscience. The ways in which economic factors like economic stability, employment security, and income disparity interact with social factors like community support and the quality of governance have been researched. Environmental conditions such as climate stability and air quality have also come to be regarded as key components of well-being [5]. Some key developments that have shaped our current understanding of global happiness:

1. The rise of positive psychology, a shift in focus from cure to cultivation [6];
2. Improved measurement tools may allow for proper cross-cultural comparisons, such as the World Happiness Report and the Better Life Index [7];
3. The growing recognition that environmental sustainability bears on long-term happiness, notably in the context of climate change [8];
4. The bearing of social inequalities on long-term happiness, by showing that more equal societies are associated with heightened well-being [9];
5. The ever-increasing influence of digital technology: changing how we connect socially and live our lives in every way possible [10];
6. An increase in emphasis on mental health as a key element in well-being [11];

The above comprehensive account showcases the multilayered inter-relationships amongst social, economic, and environmental factors determining the levels of happiness in the global context. Such interactions shall evolve as societies evolve and hence will call for an upgrade in policies and continuing orientation of research.

2 State of art

This state of art review aims to synthesize and discuss recent findings from four studies, analyzing the factors influencing happiness and well-being from different perspectives.

2.1 The economy and happiness

Though economic prosperity has been seen to benefit such things as life satisfaction, its impetus is hardly linear. Jebb et al. identified satiation points with respect to income in their investigation, implying that after reaching a point of income merit, incremental increases will slow down the resulting increase in happiness. The satiation point varies throughout the world; for wealthier countries, it is comparatively higher. Similarly, [13] found support for per capita GDP as an important predictor of life satisfaction, but individual characteristics along with country peculiarities also accounted for variation. The Easterlin Paradox supports these findings, proposing that increased income does not necessarily coincide with permanent increases in happiness. On this basis, Kuc-Czarnecka et al. discuss the Sustainable Development Goal 7 (SDG7) implementation in the European Union (EU) and its links to economic factors. Even if these writers considered that economic growth would help foster a sustainable energy transition, they showed that ecological factors play a more prominent role in the settings for switching to cleaner energy. From this thought, the authors argue, on the contrary, that one cannot alone rely on economic growth to foster the combined existence of sustainability and well-being, which comes indeed from integrating environmental considerations into a development strategy.

2.2 Social and Institutional influences on happiness

Social structure and institutional trust are great in determining general happiness, and this extends beyond merely having wealth means. Jebb et al. highlight that income satiation is not a requisite universal threshold; rather it is culturally and socially constructed, and the same cultural and social norms dictate the contribution of financial resources to well-being. Their study indicates that social cohesion and quality of governance mediate between economic prosperity and happiness. In the same context, Kuc-Czarnecka study the implementation of SDG7 in the European Union with special mention of social factors such as education and health as important deciding powers in the transition to sustainable energy and therefore affecting well-being. This creates a strong case for any proposed investment in people and human capital for the establishment of institutional health and wealth needed for sustainable long-term happiness [14]. Supporting this, Ahmadiani and colleagues propose social capital and religiosity as vital factors impacting life satisfaction in the nations studied. Their analysis reveals that these factors, along with environmental factors such as pollution, have a significant impact on happiness [13]. This implies the importance of community bonding and institutional trust as the foundational work for collective well-being. Again, this affirms the belief that it is social structures, rather than mere economic indicators, that are vital in cultivating happiness with sustainability into the future.

2.3 Environmental sustainability and well-being

The relationship between environmental factors and general well-being has gained importance from some time now. Research by Chen examined economic and environmental factors affecting life span across developed and developing nations and concluded with observable discrepancies [15]. They stated that while economic theory predicts that GDP per capita should increase life expectancy in rich countries, the opposite is expected in less developed regions owing to environmental degradation suffered by rapid industrialization. Certainly, CO2 emissions and air pollution have played their role in improving life in lower-income regions, where regulatory policies are not that stringent. This topic is supported by Kuc-Czarnecka, who argue that ecological considerations are particularly important for promoting sustainable energy practices in the EU [14]. This reflects the wider conversation on these topics, such as those viewed in the Global Happiness Report [1], which underscores environmental sustainability as key to long-term well-being. Ahmadiani added to this literature by arguing pollution to have direct influence over life satisfaction, emphasizing much need policies mitigating environmental problems for health and general quality of life [13].

2.4 The multidimensional nature of well-being

Well-being has to be looked at by engaging an encompassing approach involving economic, social, and environmental aspects. As noted by Jebb and Kuc-Czarnecka, factors like social integration and good governance contribute significantly to explaining the differences in happiness, while pure economic factors alone are insufficient[12] [14]. Likewise, Chen talk about the environmental implications of economic growth, further analyzing the necessity of implementing a policy that embodies a favorable equilibrium between development and environmental sustainability [15]. Yet another layer of argument comes from Ahmadiani who proved that, besides good economic arrangements, life satisfaction is contingent upon institutionalized trust, community networking, and environmental quality [13]. These findings take center stage in the larger debate of global happiness as they emphasize the need for an integrated approach. As old societies continue to evolve into contemporary times, the establishment of frameworks that deal with collective well-being together with continued economic stability, institutional reliability, and ecological health becomes increasingly urgent. Policy-making in the future must focus on social equity and building resilience of communities while ushering sustainable economic development in order to produce long-ranging pathways for happiness on a global scale.

2.5 Problems and goals

After a good deal of world wealth and technological enhancements, the societies nevertheless persistently fail to find sustainable happiness. The social, economic, and environmental interactions are too complex to accurately assess without deeper analyses [16]. The problems for consideration are listed below:

1. **Consciousness:** The Easterlin Paradox reveals that economic growth does not always engender increases in human happiness [17]. With acute disparities in wealth production, numerous developed countries register stagnation or decline in their well-being.
2. **Environmental degradation:** Climate change and urbanization threaten our quality of life, but the interplay of environmental factors with happiness has seldom been examined [18].
3. **Societal inequality:** Economic and social disparities have restricted access to well-being resources and hence collectively affect happiness and policy efficacy [19].

Thus, the aim of this study would be to evaluate the interaction between social, economic, and environmental factors to find more sustainable pathways to enhance well-being:

• To explore the relationship between economic indicators and happiness across societies.

• To assess the impact on well-being of environmental variables (air/water quality; green spaces climate vulnerability and policies).

• To explore social influences-trust, community support, cultural values, and governance.

• To develop a framework for the analysis of key determinants of happiness.

• To offer well-grounded recommendations for policy-making that address local, national, and global contexts.

3 Methodology

3.1 Datasets used

For a comprehensive analysis of the social, economic, and environmental factors underlying global happiness, four complementary datasets were selected. Together, these datasets provide distinct measurements through which several dimensions of human well-being reviewed in the above literature can be addressed.

1. **World Happiness Report 2024**

   Source: Kaggle (https://www.kaggle.com/datasets/ajaypalsinghlo/world-happiness-report-2024)

   Characteristics:

   • **Main variables:** Happiness ranking, happiness score, GDP per capita, social support, healthy life expectancy, freedom to make life choices, generosity, perception of corruption, positive, and negative affect;

   • **Geographic coverage:** 143 countries;

   • **Time period:** 2024 data, with possibility for comparison with previous years;

   • **Data format:** CSV (Comma-Separated Values);

   • **Data types:** Numerical (scores and metrics) and categorical (country names, regions).

   Identified problems:

   • Missing values in some metrics for certain countries, especially politically unstable or recently conflicted ones;

   • Variance within collection methodologies in different regions may affect direct comparisons;

   • Possible cultural bias in regarding the interpretation of such concepts as "happiness" and "well-being" across different cultures.

2. **Global Data: GDP, Life Expectancy and More**

   Source: Kaggle (https://www.kaggle.com/datasets/arslaan5/global-data-gdp-life-expectancy-and-more)

   Characteristics:

   • **Main variables:** GDP per capita, life expectancy, infant mortality rate, access to electricity, Human Development Index (HDI), unemployment rates, CO emissions, and internet access;

   • **Geographic coverage:** 195 countries;

   • **Time period:** Historical series with data provided across several years, making possible trend analysis;

- **Data format:** CSV (Comma-Separated Values);
- **Data types:** Numerical (socioeconomic and environmental indicators) and categorical (countries, regions, and development classifications).

Identified problems:

- Temporal inconsistency, with some variables updated more recently than others;
- Missing data concerning low-income countries or those with deficient statistical infrastructures;
- Potential methodological discrepancies in data collection between different indicators.

3. **Social Progress Index Rankings**

   Source: Kaggle (https://www.kaggle.com/datasets/samithsachidanandan/social-progress-index-rankings)

   Characteristics:

   - **Main variables:** Overall Social Progress Index score and its dimensions: Basic Human Needs (nutrition and basic medical care, water and sanitation, shelter, personal safety), Foundations of Well-being (access to basic knowledge, access to information and communication, health and wellness, environmental quality), and Opportunity (personal rights, freedom and choice, inclusiveness, access to advanced education);
   - **Geographic coverage:** 168 countries;
   - **Time period:** Most recent available data;
   - **Data format:** CSV (Comma-Separated Values);
   - **Data types:** Numerical (scores in different dimensions) and categorical (countries, regions).

   Identified problems:

   - Complexity in interpreting composite indices, which aggregate multiple dimensions;
   - Potential subjective character in certain metrics related to rights and inclusion;
   - Some dimensions may be invariably valued more heavily in the final index.

4. **Global Suicide Statistics**

   Source: Kaggle (https://www.kaggle.com/datasets/arpitsinghaiml/global-suicide-statistics?select=suicide-rate-by-country-2024.csv)

   Characteristics:

   - **Main variables:** Suicide rates by gender, age group, year, and methods, per country;
   - **Geographic coverage:** 101 countries;
   - **Time period:** Data up to 2024, with history from previous years for trend analysis;
   - **Data format:** CSV (Comma-Separated Values);
   - **Data types:** Numerical (rates and counts) and categorical (countries, age groups, gender).

   Identified problems:

   - Some countries have severe under-reporting due to cultural stigmatization attached to suicide;

• Variations in registration and classification systems across countries. • Lack of data for countries whose cause of death registration systems are not well established;

• Cultural and religious sensitivity affecting data accuracy.

3.2 Datasets integration

The combination of these four datasets will result in a multidimensional analysis that incorporates the economic, social, and environmental aspects of global well-being. Integration will be performed using countries as a common key, allowing for correlations between:

• Economic indicators (GDP per capita, unemployment) and happiness levels;

• Social factors (social support, freedom, corruption) and well-being;

• Mental health indicators (suicide rates) and their relationship with socioeconomic factors;

• Environmental impacts (environmental quality, emissions of CO2) on perceived happiness;

• Social progress and its correlation with subjective measures of well-being This integrative approach will allow testing the hypotheses raised in the literature review, particularly those related to the Easterlin Paradox and the multidimensional nature of human well-being.

3.3 Data Integration and Processing Report

This report continues the data integration and processing efforts for the global happiness analysis project. The goal remains to clean and integrate multiple datasets that capture social, economic, and environmental factors influencing happiness and well-being across countries. This work builds upon previous analyses to ensure consistency and accuracy in data preparation.

3.3.1 Datasets Description

The final integrated dataset consists of information derived from:

• World Happiness Report 2024 – Contains happiness scores, rankings, and explanatory variables such as GDP per capita, social support, and freedom.

• Global Economic and Health Data – Includes socioeconomic indicators such as GDP per capita, life expectancy, and environmental metrics.

• Social Progress Index Rankings – Provides metrics on Basic Human Needs, Foundations of Well-being, and Opportunity across countries.

• Global Suicide Statistics – Includes suicide rates by country, gender, and age group.

3.3.2 Data Processing Methodology

Step 1: Initial Data Inspection

Each dataset was examined for:

• Data structure and format

• Variable types and distributions

- Missing values and outliers
- Naming inconsistencies

Step 2: Data Cleaning

- Standardization of Country Names

o Applied a mapping function to unify country names across datasets (e.g., "United States" → "United States of America", "Czech Republic" → "Czechia").

o Trimmed whitespace inconsistencies in country names.

- Happiness Report Data

o Selected key variables relevant to happiness assessment.

o Standardized country names to match other datasets.

o Imputed missing values using regional means.

- Global Economic and Health Data

o Selected relevant economic and health indicators.

o Standardized measurement units across datasets.

o Applied log transformation to GDP per capita for normalization.

- Social Progress Index Rankings

o Standardized dimension scores to a 0-100 scale.

o Addressed missing values using domain-specific imputations.0

- Global Suicide Statistics

o Aggregated suicide rates across age groups to obtain country-level statistics.

o Normalized rates per 100,000 population.

o Adjusted under-reporting using WHO-based correction factors.

Step 3: Handling Missing Values

- Countries with **more than 50% missing values across all columns** were removed from the dataset.
- For countries with **less than 50% missing values**, missing values were imputed using the **mean of the respective factor** across similar countries or regions.

Step 4: Data Integration

Phase 1: Variable Standardization

- Created a standardized country identifier.
- Harmonized variable names and units of measurement.
- Normalized index-based variables.

Phase 2: Data Merging

- Merged datasets using an outer join to retain as much information as possible.
- Resolved conflicting values by prioritizing the most recent or reliable source.
- Applied an aggregation function to combine overlapping variables using the mean value.

Phase 3: Post-Merge Processing

- Standardized regional classifications.
- Applied feature scaling for comparability.
- Ensured consistency in data formats across variables.

3.3.3 Quality Assurance

To ensure data integrity, several validation steps were undertaken:

- **Completeness Check** – Verified country count post-merging.

- **Consistency Verification** – Cross-checked key indicators.
- **Outlier Detection** – Applied the IQR method.
- **Correlation Analysis** – Identified inconsistencies between similar variables.

3.3.4 Final Dataset Structure

The integrated dataset includes:

- **143 countries** with complete records.
- **51 variables** covering economic, social, and environmental dimensions.
- Standardized metrics for happiness, well-being, development, and health.
- Fully documented variable definitions and sources.

Key variable categories:

- Happiness and well-being metrics.
- Economic indicators.
- Social progress factors.
- Health measures.
- Environmental and governance indicators.

3.3.5 Limitations and Considerations

- Some countries have partial data due to inconsistent source coverage.
- Cultural biases in happiness measurement persist.
- Temporal inconsistencies exist across data sources.
- Under-reporting in suicide rates affects reliability for certain regions.

3.3.6 Conclusion

The data integration process successfully combined multiple sources into a unified dataset, providing a multidimensional view of happiness determinants across countries. This dataset enables further analysis of the relationships between social, economic, and environmental factors influencing well-being, ensuring consistency with prior research efforts and allowing for continued exploration of global happiness trends.

4 Data Integration and Processing: A Performance Comparison

In this section, we compare the effectiveness of two distinct methods, one traditional, using Pandas, and the other PySpark, for joining the four datasets used in the analysis of global happiness. The aim of this comparison was evaluating the computational effectiveness and potential for scaling with larger datasets.

4.1 Traditional Approach with Pandas

Pandas was used in the conventional approach to read, clean, and join the datasets. Pandas is well-known for processing data in-memory, which is especially helpful when working with small datasets. The conventional Pandas method was adequate for combining the datasets because of the integrated dataset's current size, which is roughly

150 rows and 50 columns. The join operation took 1.36 seconds to execute. This outcome demonstrates Pandas' exceptional efficiency for datasets of this size, offering a dependable and rapid method of merging data from various sources.

4.2 PySpark Approach

Since PySpark is made for processing large amounts of distributed data, it was also taken into consideration as a substitute for the join operation. However, the computational advantages of using Spark were not evident because of the comparatively small size of our integrated dataset. Furthermore, the PySpark code was unable to run properly during the initial setup due to problems with the environment configuration. Given the small size of the dataset, it is unlikely that we would have observed appreciable performance gains even if PySpark had been fully operational. Larger datasets, where distributed processing across several nodes provides noticeable advantages, are usually the domain of PySpark.

4.3 Conclusion

The findings demonstrate that Pandas was the more effective option for a dataset with this size. The traditional join method took 1.36 seconds, which is ideal for memory-fitting datasets that are small to medium in size. The small size of the dataset would have prevented PySpark from offering any appreciable performance gains in this instance, despite its strength in managing larger datasets. PySpark would be more suitable for larger datasets because it is made to take advantage of distributed computing resources, which greatly reduces processing time.

Although both approaches were theoretically feasible, Pandas was unquestionably the superior option in this particular situation. It performed quickly and simply, which made it ideal for the demands of the current project. It would be advised to use PySpark in the future to fully explore the potential of distributed processing for much larger datasets. The main lesson is that the size and complexity of the dataset, along with the available computational resources, should determine which data processing tool is used.

5 Final Dataset Storage in the Database

5.1 Databases Options

Before deciding on a database for this project, a number of options were examined. Every one of these databases has unique benefits and is better suited for certain applications. An outline of the main options taken into consideration is provided below.

- **PostgreSQL**:
  - **Importance**: PostgreSQL is a popular relational database management system (RDBMS) for projects centered around data. It is renowned for its sophisticated support for SQL operations, extensibility, and capacity to manage sizable datasets. PostgreSQL is perfect for analytical work and Big Data projects because of its strong Python integration with libraries like`psycopg2` and `SQLAlchemy`.

- **When to use**: PostgreSQL is suited for applications needing scalability, high data integrity, and complex queries. It is a great option for systems that manage substantial volumes of structured data, intricate relationships, and situations where Python integration is necessary for data analysis. However, if not set up correctly, problems like server communication issues can prevent it from being deployed.

- **MySQL**:
  - **Importance**: MySQL is a well-liked open-source RDBMS with a reputation for speed and usability. It is frequently utilized in web applications and is well suited for read-heavy tasks.
  - **When to use**: MySQL is the best choice for a quick and lightweight RDBMS for web applications with moderate data volumes. MySQL is a good option for Python-based data handling because it supports Python (via `mysql-connector-python` or `PyMySQL`). It might not be able to handle complex queries as well as some other options, though, similar to PostgreSQL.

- **SQLite**:
  - **Importance**: SQLite is a serverless, lightweight RDBMS that stores data in a single file, it is simple to use locally and deploy without requiring a server installation. For small to medium-sized projects that don't need a lot of concurrency, it is perfect.
  - **When to use**: SQLite is most suitable for projects that prioritize simplicity, speed, and ease of setup. Local testing, small applications, or situations where the dataset is manageable (like this project) are common uses for it. SQLite is a strong candidate for this project since it makes simple to receive data through Python scripts. It does not, however, scale well in high-concurrency applications or distributed environments.

- **MariaDB**:
  - **Importance**: MariaDB is a fully open-source fork of MySQL that has been performance-optimized. Improved query optimization and the Aria storage engine are two of its improved features.
  - **When to use**: MariaDB is a good substitute if MySQL isn't able to meet performance requirements. It works well for projects requiring high-performance RDBMS capabilities or web applications with moderate to high traffic. Additionally, it works well with Python thanks to connectors like `PyMySQL`. However, MariaDB's extra features might not offer much of a benefit over SQLite or MySQL for a project this small.

- **MongoDB**:
  - **Importance**: MongoDB is a NoSQL database that stores large amounts of unstructured or semi-structured data with schema-less documents that provides flexibility and scalability.

  - **When to use**: MongoDB is suited for projects requiring real-time analytics, flexible data models, or situations where data doesn't neatly fit into relational tables. It might not be the best option for structured data with a relational structure and a clear schema. It is less appropriate for this specific project because it is more difficult to query using SQL syntax, even though it integrates with Python through the `PyMongo` library.

- **Oracle Database**:
  - **Importance**: Oracle is a robust, enterprise-class RDBMS that is renowned for its sophisticated features, scalability, and security. Large datasets and intricate transactions are supported.
  - **When to use**: Oracle works best with big businesses or applications that need robust transaction support, high availability, and in-depth analytics. However, it is not a viable option for small-scale or budget-conscious projects due to its complexity and licensing fees. It can also be integrated with Python, though it might take more setup than alternatives like PostgreSQL or MySQL.

- **Microsoft SQL Server**:
  - **Importance**: SQL Server is a an all-inclusive RDBMS with robust Microsoft technology integration. It provides sophisticated features like strong security, business intelligence tools, and reporting.
  - **When to use**: SQL Server is typically used by businesses that depend on the Microsoft tech stack and require business intelligence, security, and high performance. Libraries such as `pyodbc` offer Python integration. Like Oracle, SQL Server is less appropriate for smaller projects due to its complexity and licensing fees.

- **Apache Cassandra**:
  - **Importance**: Apache Cassandra is a distributed NoSQL database created to manage enormous volumes of data across numerous servers. It is appropriate for real-time analytics and Big Data applications due to its high availability and scalability.
  - **When to use**: Cassandra is perfect for projects that need fault tolerance and high availability, particularly when working with time-series or event-driven data. Cassandra would be excessive for structured data and small datasets like the one used in this project.

5.2 Justification for the Selection of the Database

After evaluating all provided database options, **SQLite** was selected as the best database option for this project after a number of options were evaluated, especially because of the following factors:

- **Simplicity and Deployment**: SQLite is very simple to set up and use because is a serverless, self-contained database that stores data in a single file. Because it doesn't require a server installation or an internet connection, this feature is especially helpful for local testing.

- **Data Size Suitability**: The dataset for this project is well within SQLite's handling capacity. Because SQLite is made for smaller databases, it handled this amount of data with sufficient performance and there was no degradation in the database's functionality.
- **Python Integration**: SQLite easily integrates with Python, making querying, data manipulation, and storage simple. Since Python scripts were needed to insert and retrieve data, this was essential to the project.
- **No Server Dependencies**: SQLite does not require a server or complex configurations, in contrast to PostgreSQL or MySQL. The database can be used directly from the local filesystem, making it an ideal choice.

Although **PostgreSQL** was first contemplated due to its advanced features and scalability, its implementation was stalled by **communication issues** with the PostgreSQL server. For this project, **SQLite** proved to be a trustworthy, lightweight solution that provided a simple, effective database solution that was compatible with Python scripts for data insertion and querying.

Therefore, considering the size of the dataset, ease of implementation, and ease of integration with Python, **SQLite** turned out to be the most practical option for this project, even though other databases like **MySQL**, **MariaDB**, and **PostgreSQL** are well-suited for larger-scale projects.

6 Power BI Analysis and Implementation

6.1 Potential Analyses Enabled by the Merged Dataset

The final integrated dataset, which included 51 variables spanning economic, social, environmental, and health dimensions and included 143 countries, theoretically allowed for the following in-depth analyses:

- **Cross-sectional Correlations:**
  - GDP per capita, perception of corruption, freedom, generosity, social support, healthy life expectancy, and overall happiness score in comparison to the World Happiness Report variables;
  - Correlations between happiness and CO_2 emissions and air quality indicators to evaluate the effects on the environment;
- **Regional Comparisons and Clustering:**
  - Setting up countries according to continents or income levels (such as HDI tiers) in order to identify regional or development-level trends in well-being;
  - Grouping people according to the Social Progress Index's three dimensions—Basic Human Needs, Foundations of Well-Being, and Opportunity—in order to find different social performance typologies;
- **Time-Series and Trend Analysis:**
  - Historical data on life expectancy, suicide rates by nation, GDP per capita, and happiness scores to identify policy impacts or turning points over time;

- Environmental trajectories, such as air pollution and CO_2 emissions, nd associated happiness trends to analyze co-movement;

- **Equity Studies:**
  - Analyses of income inequality or suicide rates by gender within a nation in comparison to measures of overall well-being.

6.2 Analyses Actually Conducted in Power BI

We used DirectQuery to our SQLite database to create three interactive dashboards in the delivered file:

1. **Global Happiness Map:** With hover-over tooltips revealing important indicators (GDP per capita, social support, freedom, and environmental quality), the world map choropleth displays 2024 happiness scores.
2. **GDP vs. Happiness Scatterplot:** A scatter plot with slicers that allow users to change any pair of numerical variables from the combined dataset, with the x-axis defaulting to GDP per capita (log scale) and the y-axis to the happiness score.
3. **Environmental Impact Dashboard:** Slicers and line graphs showing possible lead-lag relationships that track CO_2 emissions, the air quality index, and happiness over time for user-selected nations.

Every visual relies on server-side execution: Power BI avoids direct file imports into Power BI's in-memory model by using ODBC to send SQL queries to the SQLite file that is filled by our Python ingestion pipeline.

6.3 Technical Obstacles on macOS

We encountered significant difficulties while developing Power BI Desktop on macOS:

- **Virtualization Performance:** Slow report rendering and delayed interactions were caused by running Power BI in Parallels or another virtual machine, particularly when applying filters or refreshing visuals;
- **Missing Features and Connectors:** Under emulation, some custom visuals, the QA natural-language feature, and the SQLite native connector were either unstable or unavailable;
- **Workflow Disruptions:** Debugging time was prolonged and iterative dashboard design was hampered by frequent virtual machine crashes and context switching between macOS and Windows.

These issues made the project more complicated and caused a slight delay in project milestones.

6.4 Summary

We created a reproducible, server-backed BI pipeline that strictly complies with the "no direct file import" rule by staging the entire merged dataset in SQLite and connecting via DirectQuery. The final Power BI dashboards offer high-performance, interactive insights into the multifaceted drivers of global happiness, despite the setup and development time being prolonged by macOS compatibility issues.

References

[1] Helliwell, J.F., Layard, R., Sachs, J.D., De Neve, J.-E., Aknin, L.B., Wang, S.: World happiness report 2021. Technical report, UN Sustainable Development Solutions Network (2021). https://worldhappiness.report/ed/2021/

[2] Stiglitz, S.A.F.J.P. Joseph E.: Report by the Commission on the Measurement of Economic Performance and Social Progress (2009). https://unesdoc.unesco.org/ark:/48223/pf0000206999.locale=en

[3] Easterlin, R.A.: Why Does Happiness Respond Differently to an Increase vs. Decrease in Income? (2023). https://papers.ssrn.com/abstract=3908860

[4] Joshanloo, M., Weijers, D.: A two-dimensional conceptual framework for understanding mental well-being. PLoS ONE **14**(3) (2019) https://doi.org/10.1371/JOURNAL.PONE.0214045

[5] Diener, E., Seligman, M.E.P.: Beyond money: Progress on an economy of well-being. Perspectives on Psychological Science **13**(2), 171–175 (2018) https://doi.org/10.1177/1745691616689467

[6] Seligman, M.E., Csikszentmihalyi, M.: Positive psychology. an introduction. The American Psychologist **55**(1), 5–14 (2000) https://doi.org/10.1037/0003-066X.55.1.5

[7] Economic Co-operation, O., Development: How's Life? 2020 (2020). https://doi.org/10.1787/9870C393-EN

[8] O'Brien, C.: Sustainable happiness: How happiness studies can contribute to a more sustainable future. Canadian Psychology **49**(4), 289–295 (2008) https://doi.org/10.1037/A0013235

[9] Walton, E.: The inner level: How more equal societies reduce stress, restore sanity and improve everyone's well-being: Inequality as it affects individuals. The British Journal of General Practice **69**(686), 452 (2019) https://doi.org/10.3399/BJGP19X705377

[10] Twenge, J.M.: More time on technology, less happiness? associations between digital-media use and psychological well-being. Current Directions in Psychological Science **28**(4), 372–379 (2019) https://doi.org/10.1177/0963721419838244

[11] Keyes, C.L.M., Martin, C.C.: The complete state model of mental health, pp. 86–98. Cambridge University Press, Cambridge, United Kingdom (2017). https://doi.org/10.1017/9781316339275.009

[12] Jebb, A.T., Tay, L., Diener, E., Oishi, S.: Happiness, income satiation and turning points around the world. Nature Human Behaviour **2**(1), 33–38 (2018) https://doi.org/10.1038/s41562-017-0277-0

[13] Ahmadiani, M., Ferreira, S., Kessler, J.: What makes people happy? evidence from international data. Journal of Happiness Studies **23**(5), 2083–2111 (2022) https://doi.org/10.1007/s10902-021-00478-y

[14] Kuc-Czarnecka, M., Markowicz, I., Sompolska-Rzechuła, A., Stundziene, A.: Assessing sustainable development goal 7 implementation and its nexus with social, economic and ecological factors in eu countries. Sustainable Development (2024) https://doi.org/10.1002/sd.3156

[15] Chen, Z., Ma, Y., Hua, J., Wang, Y., Guo, H.: Impacts from economic development and environmental factors on life expectancy: A comparative study based on data from both developed and developing countries from 2004 to 2016. International Journal of Environmental Research and Public Health **18**(16) (2021) https://doi.org/10.3390/ijerph18168559

[16] Simonova, M.D.: Beyond gdp. measuring what counts for economic and social performance. MGIMO Review of International Relations **12**(5), 172–180 (2019) https://doi.org/10.24833/2071-8160-2019-5-68-172-180

[17] Powdthavee, N., Ward, G., Flèche, S., Layard, R.: The Origins of Happiness, (2018). https://doi.org/10.1515/9781400889129

[18] Bernauer, T., McGrath, L.F.: Simple reframing unlikely to boost public support for climate policy. Nature Climate Change **6**(7), 680–683 (2016) https://doi.org/10.1038/NCLIMATE2948

[19] Oishi, S., Kesebir, S., Diener, E.: Income inequality and happiness. Psychological Science **22**(9), 1095–1100 (2011) https://doi.org/10.1177/0956797611417262