Business Intelligence: Applications & Projects

GlobalBite: A Data-Driven Analysis for Restaurant Expansion in Major Cities Worldwide

Group 1
JinYoung Jeon
Mohammad Azizul Kawser
Nandor Hajdu
Phung Tran
Trang Nguyen

Contents

1. Introduction

1.1 About DashDelicious

DashDelicious (DD) is a fast-food chain that has become a household name in the United States since its establishment in 2015. The company has over 2000 stores covering both the Northern and Southern US areas and a workforce of 3500 to 4000 employees. DD's success can be attributed to its innovative product portfolio, ability to adapt to changes in market conditions, and reasonably competitive pricing strategy.

DD's strategic focus is on expanding its market share both within the United States and internationally. The company aims to capitalize on its brand recognition and accomplishments in the US market to penetrate other countries. DD achieves this goal by consistently innovating and introducing new products to satisfy its customers' changing preferences in various regions.

1.2 Factors of Success

- **Adapting to Local Tastes and Cultures:** One of the primary reasons for DD's success is its ability to adapt to local tastes and cultures. The company recognizes that different regions and demographics have unique food preferences, and it has tailored its product offering accordingly. DD is also dedicated to offering healthy food options on its menu, recognizing the increasing importance of health consciousness among customers.
- **Reasonably Competitive Pricing Strategy:** The company offers affordable prices for its menu items, making it an attractive option for customers who are looking for budget-friendly fast-food. DD's pricing strategy has been instrumental in maintaining its market share in the US market. The company plans to replicate this pricing strategy in its international expansion efforts to remain competitive in new markets.
- **Data-Driven Decision-Making:** DD uses advanced analytics and data science to gain insights into consumer preferences, market trends, and operational efficiencies. This enables DD to make informed decisions regarding product development, pricing, and expansion strategies, resulting in a competitive advantage over its rivals. By leveraging Business Intelligence to optimize its operations, DD can deliver value to its customers while maintaining a profitable business model.
- **Strong Financial Performance:** DD has consistently reported solid financial results, including robust revenue growth and profitability. DD's ability to generate strong financial returns has enabled the company to reinvest in the business, expand into new markets, and innovate its product offerings. DD's financial strength has played a critical role in its success and ability to sustain long-term growth.

1.3 GlobalBite Project

DD understands the importance of market research and data analysis in identifying new growth opportunities. The company has invested heavily in building a robust Business Intelligence team to inform its expansion strategy. By leveraging data, DD can identify new markets that offer favorable conditions for expansion, such as a growing middle class, high disposable income, and a low cost of living.

GlobalBite is a data-driven analysis project that aims to identify potential new markets for DD. The project is part of a broader market research campaign aimed at expanding DD's presence in major cities worldwide. GlobalBite focuses on the cost of living, expense, and salary data to pinpoint potential cities for deeper research into potential new markets for DD.

Through GlobalBite, DD can better understand the market conditions in different cities, allowing the company to make more informed decisions about where to focus its expansion efforts. This data-driven method enables DD to stay ahead of its competition by being proactive in identifying potential new markets.

1.4 Business Questions

GlobalBite is designed to answer several critical business questions related to potential new markets:

- The project focuses on comparing the cost of living in various regions. This will involve determining the average cost of living index worldwide and how this varies by region or country in which DD targets to operate. This information is essential for identifying cities that offer cost-effective expansion opportunities for DD.
- The project aims to explore potential customers' purchasing power in different cities worldwide. This will involve analyzing the average salary in different cities and the cost of different factors such as rent, groceries, or restaurant prices. The project will also examine which expenses have the most significant impact on the cost of living in different cities. Understanding these factors is crucial in determining the affordability of DD products and services in potential new markets.
- The project will investigate the difference in the cost of space between large metropolitan areas and smaller cities. This is important in identifying new markets offering a cost-effective expansion opportunity for DD. By understanding the cost of space in different cities, DD can identify new market opportunities that offer optimal returns on investment.
- The project aims to develop a scoring model for cities around the world based on the available data. This scoring model will consider factors such as cost of living, purchasing power, and other relevant metrics to determine the most favorable cities for DD's potential expansion. This will enable the company to make informed decisions about which cities to target, based on their potential for profitability and success.

1.5 Data Source

During the data exploration phase, the Business Intelligence team discovered a valuable open dataset from Kaggle.com, which contained an extensive collection of data on the cost of living in various cities around the world. The team recognized the potential usefulness of this data source, as it could provide us with valuable insights into the economic conditions in different regions and cities.

As we began to dive deeper into the dataset, the team realized that there were a few missing pieces of information that were necessary to conduct a more comprehensive analysis. For example, while the cost-of-living data was incredibly valuable, it would be difficult to draw meaningful conclusions without additional factors such as population.

To overcome the data deficiencies, the Business Intelligence team decided to combine the primary dataset with additional relevant datasets to gather the required information. The ultimate dataset was obtained by merging three separate datasets, specifically the “Global Cost of Living”, “Cost of Living Index 2022”, and “Geonames All City with a Population”. Details regarding the final dataset and the methods employed to analyze it will be presented in the forthcoming sections.

2. Project Approach and Methodology

2.1 Business Framework for Expansion

As discussed above, there are several considerations that the GlobalBite project aims to address. To accomplish this, Business Intelligence (BI) follows the market expansion mode (MEM) model by Robinson & Lundstrom (2003) as shown in Figure 1. The MEM model is a comprehensive framework for evaluating market expansion opportunities based on the alignment between internal and external factors presented by the first model. This process involves assessing the degree to which the organization aligns with these factors and selecting an appropriate strategy (such as direct ownership, exporting, joint venturing/licensing/contract management) based on factors like risk, resource requirements, and desired level of control (Robinson & Lundstrom, 2003, p. 268).

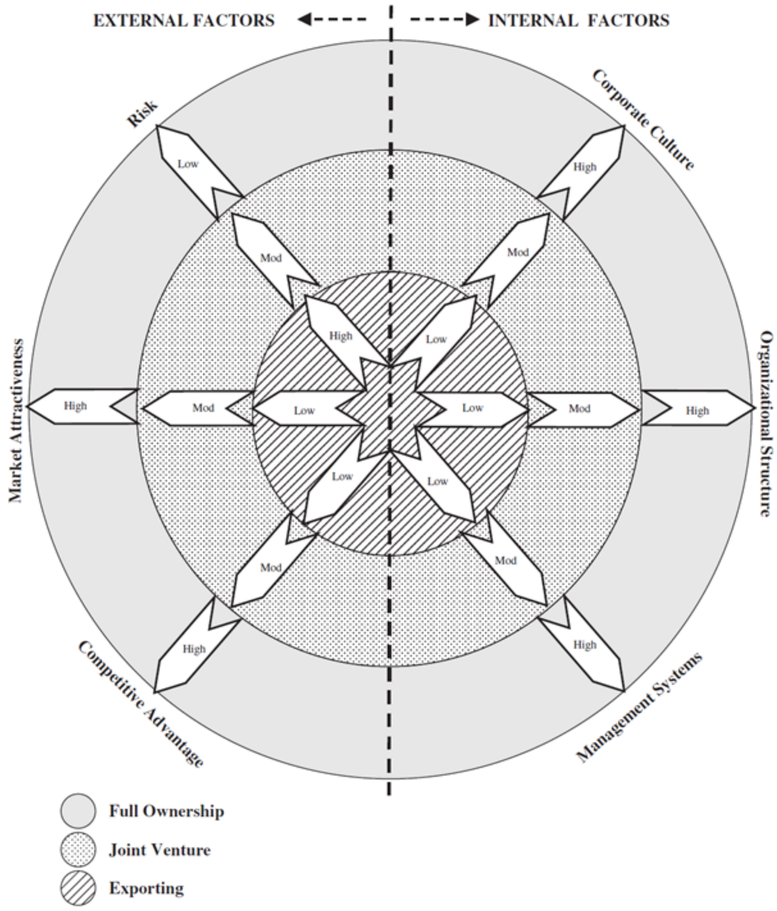

Figure 1. Market expansion mode (MEM) model (Adapted from Robinson & Lundstrom, 2003, p. 265)

The current focus of the BI team is to assess the market attractiveness, which is defined as the potential appeal of a market that an organization is contemplating entering is influenced by several factors such as the labor force demographics, population characteristics, and cultural norms (Robinson & Lundstrom, 2003, p. 265). For example, a market that is deemed desirable for an antique store may consist of a larger percentage of affluent and upper-class individuals. Similarly in our case, a city where there is a significant eating-out culture, and meals are affordable compared to cost of living factors, we can assume profitable operation in that location. We are considering franchising for international operations, which is further discussed later.

Another framework we use in our research into expansion strategy is the OLI paradigm, also referred to as the eclectic paradigm. The OLI paradigm serves as a unifying framework for understanding the degree and pattern of foreign-owned activities, which arise from three sets of advantages: Ownership, Location, and Internalization (OLI) (Sharmiladevi, 2017). The presence or absence of OLI advantages either stimulates or hinders firms from embarking on foreign activities. It has been a well-tested and researched model in various contexts, including in the restaurant industry (Rivas, 2011).

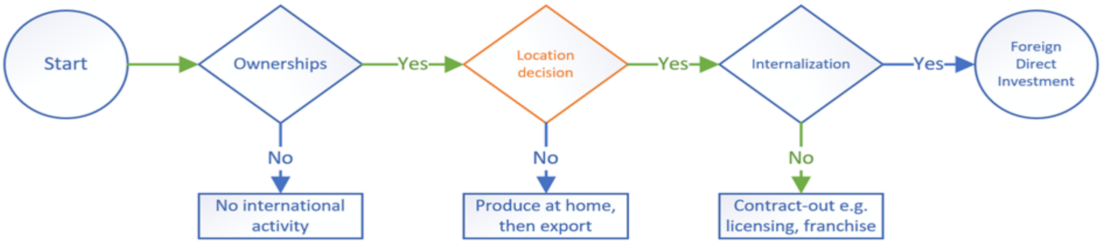

Figure 2. OLI paradigm and the factor under examination by the BI team (Welch et al., 2007)

Describing the paradigm according to Welch et al. (2007, p. 31):

- Ownerships: The ownership factor assesses whether a firm has assets that provide it with a competitive edge over local companies in a foreign market. This advantage is crucial for the firm to compete with indigenous firms due to the added expenses of operating in a foreign context, and without it, the firm may not be able to survive in a fiercely competitive foreign market.
  As mentioned before, an innovative product portfolio, the ability to adapt to market conditions, and a reasonable pricing strategy, have all played a significant role in DD's success domestically, being able to compete with major competitors. Therefore, it is plausible the success would continue internationally as well.

- Location: Originally, the location factor is concerned with whether a firm's assets are better utilized in foreign locations rather than in its home country. It determines whether foreign production is necessary and whether to service a foreign market through exports or local production. Rivas (2011) considered this factor from the perspective of the cultural and development distances between the home country and the country where the companies decided to enter in the context of the multi-nationalization of restaurants.
  In our case, we interpreted this factor to determine the profitability and sustainability of the location based on certain variables related to the cost of living. Besides, a country with a high cost of living may require higher wages for workers, leading to higher production costs, which can impact a firm's decision to operate in that city/country.

- Internalization: The internalization factor in the OLI paradigm considers whether it's best to internally transfer assets involved in an activity, and a positive answer to this question is essential for conducting activities in-house.
  Franchising is a popular method of international expansion in the restaurant industry. As defined by Vaughn (1974), it involves a parent company allowing a smaller company or individual to operate under specific conditions for a fee, including selling the parent company's products, using its name, and replicating its methods and trademarks. Direct franchising can also serve as a way to test foreign markets before making large investments, but it is most effective in countries that are close geographically and culturally (Welch et al., 2007). For the initial

international expansion of DD, direct franchising appears to be the most suitable course of action.

To summarize this section, the objective of business intelligence (BI) is to compile a list of viable locations for international expansion (franchising) and conduct a comprehensive market analysis of the selected cities. This analysis will factor in both external and internal components of the MEM and result in the selection of a single city for expansion, with the possibility of further expansion based on insights gained from the initial venture.

2.2 Approach to Data Analysis

Business intelligence aims to assist stakeholders in comprehending their organization's operations, facilitating faster decision-making, and managing operational performance, thereby allowing users to recognize emerging trends, take action, and manage organizational issues (Dedić & Stanier, 2017, p.2). In today’s knowledge-driven economy, organizations rely on BI tools to gather, evaluate, and distribute information, empowering knowledge workers to make well-informed decisions amidst stiff competition (Hedgebeth, 2007, p. 414).

To fulfill our objective of selecting suitable locations for expansion, and propose answers to the business questions we formulated, the BI team decided to apply the KDD process. KDD refers to the entire process of discovering valuable knowledge from data, of which data mining is a specific step that involves using algorithms to extract patterns from data. The other steps in the KDD process, such as data preparation, cleaning, selection, and interpretation, along with the incorporation of relevant prior knowledge, are essential in ensuring that valuable knowledge is obtained from the data (Fayyad et al., 1996, p. 28).

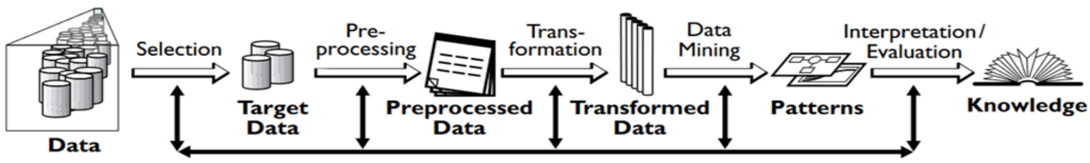

Figure 3. KDD (Knowledge Discovery in Databases) Process (Adapted from Fayyad et al., 1996, p. 29)

- Data Selection: Data selection is a process in data analysis that involves identifying and choosing relevant data for further analysis (Strengholt, 2020).
  Our team is decided to use Cost of Living Data with Indexes because it provides insight into the expenses that a company would incur when operating in a particular location, as well as the cost of goods and services to inform business decisions. Furthermore, we incorporated population and geolocation data of various cities, and subsequently merged the

datasets, ultimately resulting in a consolidated dataset. Short descriptions of the datasets:

- Global Cost of Living: the dataset compiles data on living expenses for nearly 5,000 cities worldwide, covering aspects such as housing, food, transportation, and healthcare. This information helps compare and analyze living costs across cities and countries, as well as identify trends and patterns over time.
- Cost of Living Indices 2022: offers insights into expenses related to groceries, restaurants, transportation, and utilities in approx. 500 cities. As a relative indicator, it uses New York as a reference point, making it useful for researchers and analysts exploring socio-economic matters.
- Opendatasoft Geonames - All Cities with a population > 1000: a comprehensive collection of information on cities and towns around the world with a population greater than 1,000. Derived from the Geonames geographical database, this dataset provides valuable insights into various aspects of these urban areas such as population, geolocation, and time zone data.

The chosen datasets are from Kaggle.com, under CC0 1.0 Universal (CC0 1.0) Public Domain Dedication license, which means we can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission (Creative Commons — CC0 1.0 Universal, n.d.).

- Data Cleansing: Data cleansing is a process that involves removing anomalies from existing data to obtain an accurate and consistent representation of the data in the real world. This process includes eliminating errors, resolving inconsistencies, and converting the data into a standardized format (Ridzuan & Wan Zainon, 2019). For example, we used multiple imputation methods for missing values (iterative imputation from scikit learn).

- Data Transformation: Data transformation involves modifying the selected data to ensure compatibility with the structure of the target database. This process can include changing the format, structure, semantics, or context of the data, as well as removing duplicates and reordering the data (Strengholt, 2020). After merging datasets, we performed data normalization, column renaming, and data reformatting to match a number of rows.

- Data Mining: This process involves employing statistical, mathematical, artificial intelligence, and machine learning methods to extract valuable information and knowledge from vast databases (Sharda et al., 2019).
In our analysis, we did not use any sophisticated algorithm (i.e. Machine Learning), but we performed normalization (min-max), statistical analysis such as correlation, and used mathematical formulas to create additional fields for visualization.

- Knowledge: Anything that has been learned, perceived, discovered, inferred, or understood; the ability to use information (Sharda et al., 2019). We used data visualization techniques are used to present complex data in a way that makes it easy to comprehend and draw insights from.

> Effective data visualization can help to transform raw data into meaningful information, allowing users to gain a deeper understanding of the data and extract valuable knowledge (Kumar & Belwal, 2017).

One of the crucial methods we had to use in the visualization phase is the min-max normalization technique as previously stated, for ensuring that data is comparable, accurate and interpretable.

This method involves shifting and rescaling values in such a way that they fall within the range of 0 to 1. This is achieved by subtracting the minimum value from each value and then dividing the result by the difference between the maximum and minimum values (Géron, 2019, p. 109). Although min-max normalization ensures that the original data values maintain their relationships with each other, the downside of restricting the range of values to a specific interval is that the resulting standard deviations are reduced. This can potentially dampen the impact of outliers (Ciaburro, 2018, p.291).

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Figure 4. Min-max normalization formula (Adapted from Ciaburro, 2018, p.291)

To achieve this in Tableau, we have to create a so-called Level of Detail expression (also known as LOD expression) (Tableau, 2023):

```
(x -{FIXED : MIN(x)})

/

({FIXED  : MAX(x)}-{FIXED  : MIN(x)})
```

We can replace the variable "x" with the fields that require normalization.

As an outcome of normalizing the data, we can make sure that the visual representation accurately reflects the true relationship between the variables being compared. This can lead to more accurate insights and better decision-making based on the visualization.

In conclusion, throughout our data analysis, we used Python to clean and transform the data by eliminating anomalies, fixing inconsistencies, and changing it into a standardized format for easy analysis. We then utilized Tableau to generate insightful visualizations, including interactive dashboards, to effectively communicate the findings and make the information more accessible and actionable for decision-makers. The combination of Python and Tableau allowed us to streamline the data analysis process and present the information in a way that made it easier to understand and interpret.

3. Analysis and Findings

3.1 Data Description

The Full data analyzed for this exploratory analysis comprises three different datasets found in the public domain. In the "Global Cost of Living" dataset from Kaggle, information such as 'Meal price', 'McMeal at McDonald's price', space rental cost, 'Average Monthly Net Salary (After Tax) (USD)', along with other 41 entities were extracted. The "Cost of Living Index by Cities" dataset from Kaggle provided data on the "Cost of Living Plus Rent Index", "Rent Index", "Groceries Index", "Restaurants Index", and "Local Purchasing Power Index". The third dataset, "Geonames - All Cities with a population > 1000", from Opendatasoft, included information such as Geoname ID, Population, and Coordinates. These three datasets were consolidated into one dataset, encompassing 536 cities from 115 different countries. The analysis was based on this comprehensive dataset, which provided detailed and relevant information on various cost of living factors, allowing for a thorough exploration of potential markets for DD's new restaurant venture.

3.2 Calculated Fields

Calculated fields allow us to create new data from data that already exists in the data source. When we create a calculated field, we are essentially creating a new field (or column) in our data set, the values or members of which are determined by a calculation that we can control. This new calculated field is saved to our data set in Tableau and can be used to create more robust visualizations.

In this project, we developed three calculated fields, “Affordability of Fast-Food Meals (Net Income/McMeals prices)” and “Tendency to eat out Index (GI/RPI)” to analyze and compare different factors related to affordability, consumer behavior, purchasing power and market size to identify potential markets for the DD's new restaurant venture.

- **Affordability of Fast-Food Meals (Net Income/McMeals prices):** This calculated field is derived by dividing the net income of customers by the average price of McMeals, which is considered a similar product to our offerings. This field provides insights into the affordability of fast-food meals for potential customers, with a higher value indicating better affordability. This information can help us assess the purchasing power of the target market and identify regions where customers may have higher disposable income to spend on dining out.
- **Tendency to eat out Index (GI/RPI):** This calculated field is obtained by comparing the prices of groceries and restaurants, using the Groceries Index (GI) and Restaurants Index (RPI) from the dataset. A higher value in this field indicates a greater tendency of customers to eat out at restaurants, which can provide insights into consumer behavior and preferences for dining options. This information can help us identify

regions where customers are more inclined to dine out, indicating a potential market with higher demand for restaurant services.

These calculated fields can help answer the business question of finding potential markets to enter by providing quantitative measures that allow for comparisons and evaluations of different regions based on key factors related to affordability, consumer behavior, purchasing power, and market size. By using these calculated fields in data visualizations and analysis, we can identify regions that show favorable values in these fields, indicating potential markets with higher potential for success. This can help DD to make informed decisions on where to focus its efforts in expanding its restaurant operations and tapping into new markets with growth opportunities.

3.3 Scoring Model

A scoring system is required in this context to objectively rank and compare different cities based on multiple factors that are relevant to the DD's decision-making process. The scoring system allows us for a systematic and quantitative evaluation of cities, taking into consideration various data fields that reflect important aspects such as affordability, market size, purchasing power, and customer tendencies. Before calculating the final scores, all variables are normalized to ensure they are measured on a notionally common scale. By assigning weights set by DD to each data field based on their relative importance, the scoring system ensures that different factors are considered proportionally in the evaluation process. This helps to avoid subjective biases and provides a more robust and consistent approach to comparing cities. The scoring system also facilitates decision-making by providing a clear and concise representation of city rankings. This allows DD to easily identify cities with higher scores, indicating their potential as favorable markets for the new restaurant venture. It helps to prioritize cities and make informed decisions on where to potentially enter the market based on a quantifiable evaluation.

We developed a scoring model for ranking cities that takes into consideration four calculated fields: "Affordability of Fast-Food Meals (Net Income/McMeals prices)", "Tendency to eat out Index (GI/RPI)", "Population (normalized)", and "Local Purchasing Power". Each of these calculated fields contributes to the overall score with a specific weightage assigned to it, which is based on its relative importance in the decision-making process. The scoring model assigns a weightage of 25% to "Affordability of Fast-Food Meals", which is calculated as the ratio of net income to the cost of McMeals, with a higher score indicating greater affordability for potential customers. The "Tendency to eat out Index" is assigned a weightage of 30% and is calculated as the ratio of the Groceries Index to the Restaurants Index, with a higher score indicating a greater tendency of customers to eat out at restaurants. The "Population (normalized)" is assigned a weightage of 20% and is calculated as the normalized population of a city, which represents the market size. A higher score indicates a larger potential customer base in terms of population. Lastly, "Local Purchasing Power" is assigned a weightage of 25% and represents the purchasing power of consumers in a city, with a higher score indicating a higher degree of purchasing power.

3.4 Visual Analysis

The visualization analysis part of the report focuses on presenting the visual discovery of the exploratory data analysis using various data visualizations. The visualizations are designed to provide insights into different factors such as the affordability of fast-food meals, tendency to eat out index, local purchasing power, and population which were used to develop the unified scoring system.

3.4.1 Map Plot

A map plot is created to visualize the population and cost of living data for different cities. The map plot used a geographic map as the backdrop, with bubbles overlaid on top of the map to represent cities. The size of the bubble indicated the population of each city, with larger bubbles representing higher-population cities and smaller bubbles representing lower-population cities. The size of the bubbles was scaled proportionally to the population data, allowing for a visual comparison of the population sizes across different cities. The color of the bubble indicated the cost of living in each city. A color scale was used to represent different levels of cost of living, with blue colors representing lower cost of living and red colors representing higher cost of living. The color of the bubbles was determined based on the cost of living data for each city, allowing for a visual comparison of the cost of living levels across different cities.

The map plot allowed for a visual exploration of the relationship between population and cost of living in different cities. Cities with larger bubbles and darker colors represented cities with higher populations and higher costs of living, while cities with smaller bubbles and lighter colors represented cities with lower populations and lower costs of living. The map plot is interactive, allowing users to hover over the bubbles to view detailed information about each city name, country name, population size and cost of living index.

The map plot provided a visual overview of the distribution of population and cost of living, helping to identify potential markets with larger populations and lower costs of living, which could be favorable for the new restaurant venture.

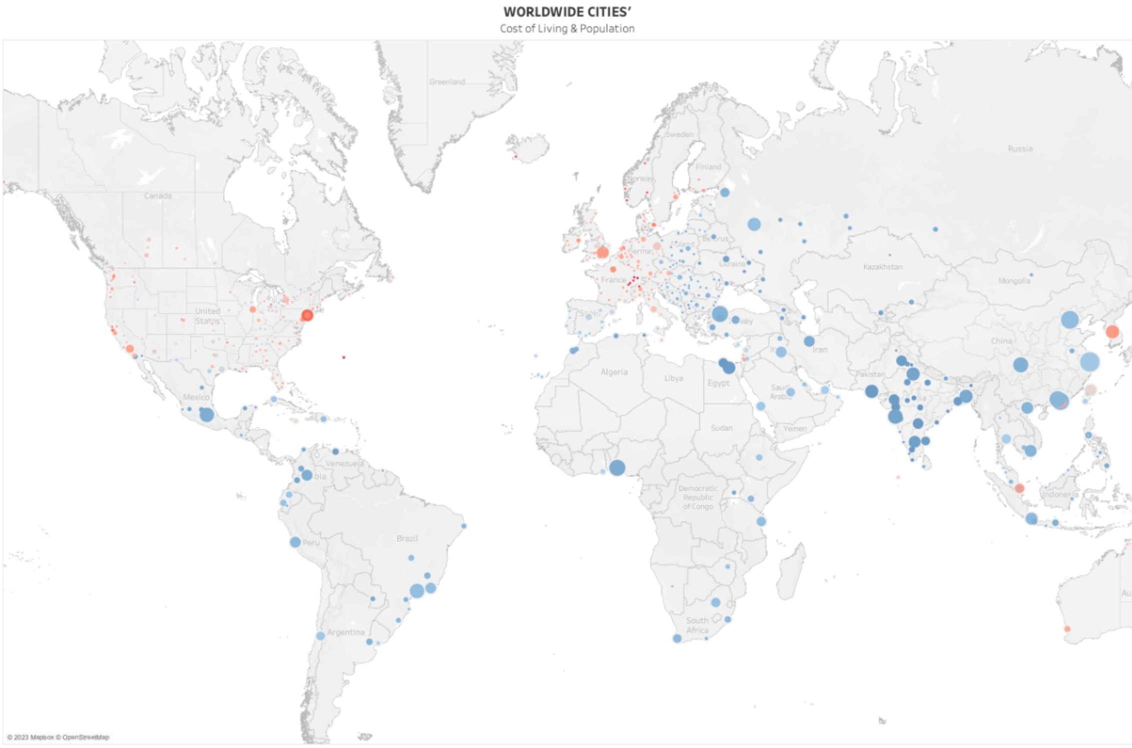

Figure 5. Map Plot visualizes the Cost of Living and Population data of cities around the world

3.4.2 Tree Map

A tree map was generated to compare the average net salary by countries and cities. In this tree map, each country was represented by a rectangle, with the size of the rectangle proportional to the average net salary after tax (USD) of that country. Countries with higher average net salary after tax were represented with larger rectangles while those with lower average net salary after tax were indicated with smaller bubbles. The rectangles were also color-coded to provide a more visualized way to compare the average net salary among countries and cities, with red colors presenting higher salaries and blue colors representing lower salaries.

The tree map was divided into two levels: country and city levels. Users can easily expand from country to city-level view and the interactive chart allows its users to hover over rectangles to obtain the average net salary information.

Tree map was a powerful tool to visually display average net salaries by countries as it allowed users to briefly spot trends and patterns regarding average net salaries, compare to their historical data and make informed business decisions. For example, a business may notice that some specific regions or countries have higher or lower average net salaries than expected. This information can be utilized by management levels to decide where to target their marketing and business development efforts or allocate further business investments and resources.

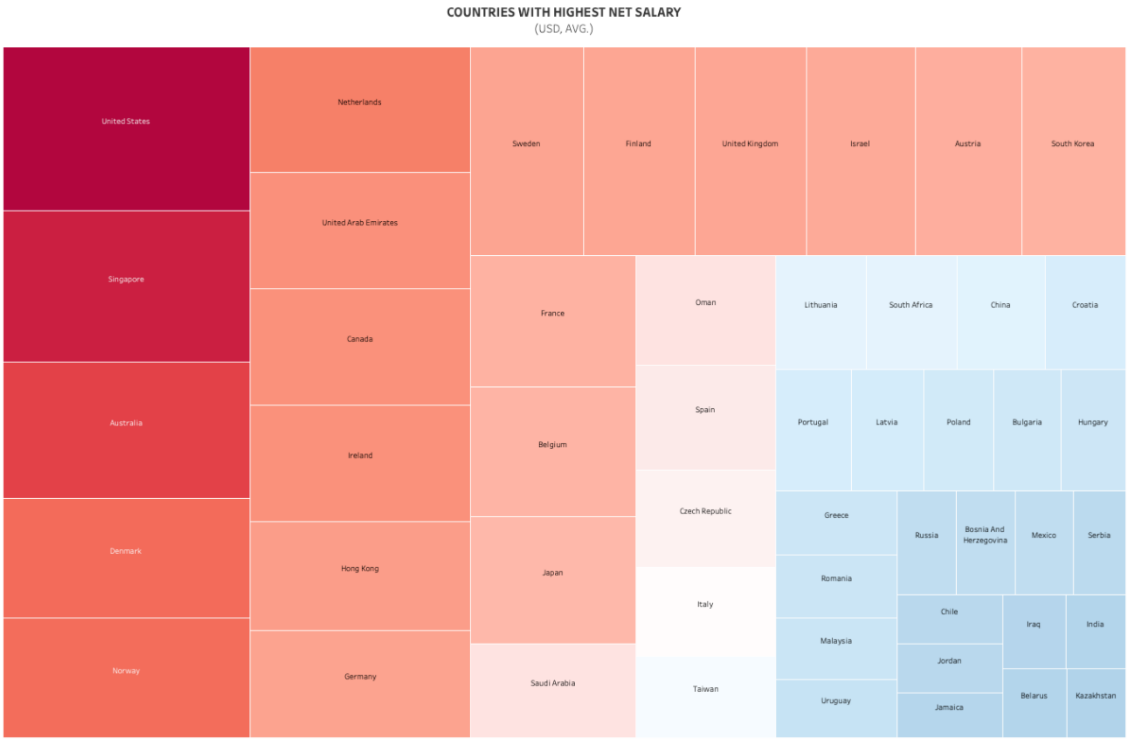

Figure 6. Treemap visualizes Average Net Salary by country

3.4.3 Scatter Plot

A scatter plot was applied to visualize and compare how the average monthly net salary and rent ratio varies across countries and cities. In this scatter plot, each data point represented a single country or city and was plotted based on its values for the two variables, with the x-axis representing the average monthly net salary variable and the y-axis representing the rent index variable. Each data point was represented by a circle and a color scale was utilized to present different levels of the ratio for the circles. The scale was defined with three levels: Comparable (blue), High rent (red) and Low rent (green), which relatively indicate which countries and cities have comparable, higher or lower rent prices.

The scatter plots also presented different quadrants indicating high salary/high rent, high salary/low rent, low salary/high rent and low salary/low rent. It can be observed that most developed countries fall into the high salary/high rent quadrant while several developing countries are scattered in the low salary/low rent quadrant.

The scatter plot was an effective tool for presenting the average monthly net salary and rent index as it enabled the business to gain some initial insights on potential rent expenses or customer purchasing power for future business expansions.

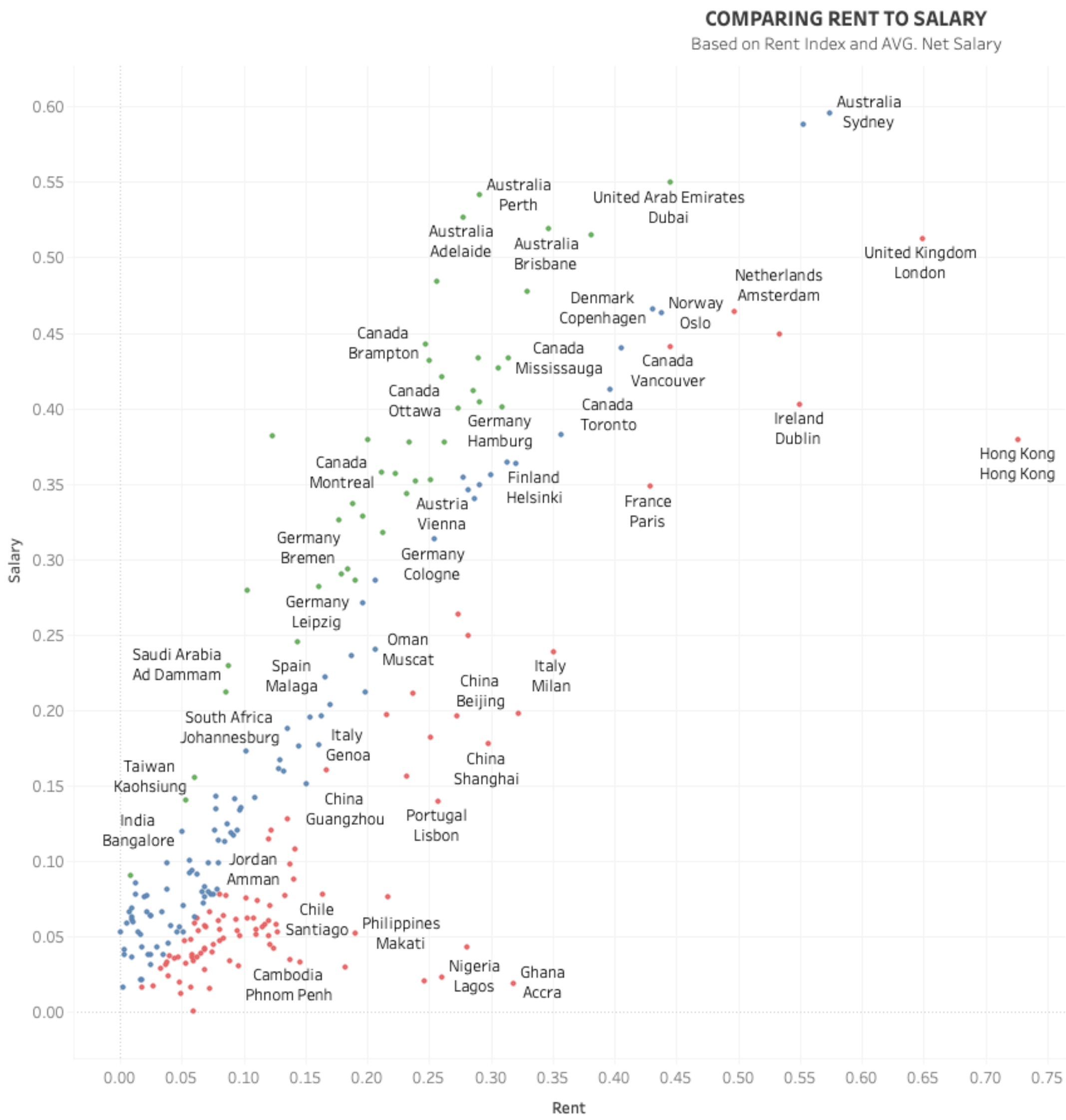

Figure 7. The scatter plot visualizes the comparability of Rent and Net Salary

3.4.4 Bar Chart

Bar charts are employed to compare the calculated scores across different countries and cities in the analysis. Horizontal bars were used to represent the scores, with each bar representing a different country or city, sorted by rank. The bar charts provided a clear visual representation of the scores, allowing for a quick and easy comparison of the rankings. The horizontal bars in the bar charts were scaled to represent the calculated scores, with longer bars indicating higher scores and shorter bars indicating lower scores.

The bar charts provided an effective way to visualize the relative performance of different countries or cities based on the calculated scores. By comparing the lengths of the bars, it was possible to quickly identify the countries or cities with the highest or lowest scores, and understand the relative rankings of different locations. The bar charts facilitated a visual exploration of the performance of different countries or cities, helping to identify potential markets with higher scores that may be more favorable for the new restaurant venture.

The bar charts were interactive, allowing users to hover over the bars to view detailed information about the scores, such as the specific values or percentages. Users could also customize the appearance of the bar charts, such as changing the colors or sorting order, to further analyze and interpret the data. The bar charts were a useful visual tool for comparing the calculated scores across different countries or cities, providing insights and supporting decision-making in identifying potential markets for the new restaurant venture.

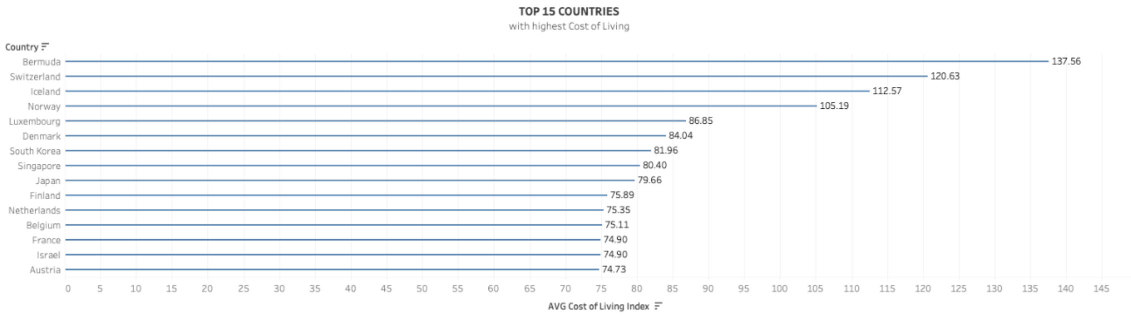

Figure 8. The bar chart shows the top 15 cities with the highest Cost of Living index

3.5 Findings and Interpretations

Based on the metrics, scoring models and visualized dashboards, some initial insights regarding market expansion are sketched for our consideration:

Potential markets for new DD businesses would be highlighted based on the average score of four defined metrics: Affordability of Fast-Food Meals, Tendency to eat out Index (GI/RPI), Local Purchasing Power and Population.

Affordability of Fast-food Meals provides insight into the cost of dining out in different countries. Countries with lower fast food meal affordability rates may indicate that consumers are more price-sensitive when it comes to dining out. However, it is also important to note that this metric may not be representative of the entire restaurant industry, as fast food represents a specific segment of the market.

In addition, the Tendency to eat out Index compares the Grocery price and the Restaurant price, with a higher index indicating a greater tendency of people to eat out rather than eating at home. Countries with a higher tendency to eat out might imply a larger potential market for the restaurant business. However, other factors that might impact the customer’s habit of eating out, such as cultural norms, availability of restaurants or dining experience should be considered in the future.

Local Purchasing Power provides insights into the relative purchasing power of consumers in different countries. Countries with higher local purchasing power may indicate that consumers have more disposable income to spend on dining out. It is also crucial to take into account other aspects, such as the cost of living and the availability of restaurants, as these can also impact consumer behavior.

Lastly, population metric can help businesses to estimate the potential size of the market for restaurant businesses. Countries with larger populations may indicate a larger potential customer base, although it is important to consider other factors, such as the level of competition and consumer preferences, as these can also impact the viability of expanding in a particular market.

From a regional perspective, the Asia Pacific region (specifically Southern Asia, Southeast Asia, Middle East and Australia) is the most potential location for the restaurant business, with several countries obtaining high rankings, whereas from a country perspective, South Korea, Taiwan, Singapore, Japan and Hong Kong are the five countries having more scope for future market study. Based on these results, more in-depth research can be conducted to analyze the gaps in these potential markets compared to our current study and identify potential business opportunities and constraints.

4. Conclusions

4.1 Summary

To conclude, our data-driven analysis conducted as GlobalBite for Dash Delicious has provided multiple insights into the market potential of major cities worldwide which will potentially aid Dash Delicious into focusing on their market research and their chain brand expansion in those locations in the future.

To answer the business questions mentioned, we have explored the market expansion opportunities for GlobalBite. To achieve this, we have utilized the Market Expansion Mode (MEM) model by Robinson & Lundstrom (2003), which is a comprehensive framework for evaluating market expansion opportunities based on the organization's alignment with internal and external factors, such as risk, resource requirements, and desired level of control. Our focus has been on assessing the market attractiveness of potential locations that will yield profitable operations for expansion and continuation of market research. Which was aligning our factors with culture and cities with significant eating-out culture, and affordable means in comparison with living factors.

Based on the analysis of the four key metrics, which are a)Affordability of Fast food meals, b)Tendency to eat out index (GI/RPI), c)Local purchasing power, and d)population, we found that countries with lower fast food meal affordability rate and higher tendency to eat out index may indicate a larger potential market for DD. Countries with higher local purchasing power and larger population also might indicate a larger potential customer base than those countries with lower purchasing power and lower population.

From a regional perspective, the Asia Pacific region was identified as the most potential location for the restaurant businesses areas South Korea, Taiwan, Singapore, Japan, and Hong Kong were countries (and cities) that have the potential for further study to find out the business success potential of DD.

With the results above, GlobalBite has successfully addressed the business question of comparing the cost of living in various regions for cost-effective expansion opportunities, as well as exploring potential customers' purchasing power around the globe. GlobalBite has successfully developed a grading scale for cities around the world in the visual analysis that DD can utilize to aid their decision on market research which will be explained further in the next section.

4.2 Recommendations for Future Business Decisions

GlobalBite project created two interactive and dynamic visualizations dashboards in the Tableau platform that will aid DD to gain deeper insights into their data and make more informed decisions. It will enable DD to quickly and easily identify trends and patterns of each region and cities in the data, as well as to present data in a visually engaging and interactive way, aiding the decision makers of DD to better understand and retain information, which can lead to more effective decision making. In short, the two dashboards created in Tableau will provide valuable insights in terms of potential markets for expansion.

Dashboard 1 focuses on the general Cost Of Living (COL) around the world. It highlights that US cities' COL indices are higher than the world average, which could make Europe a comparable market for DD if it wants to expand to a market with similar properties. Additionally, the dashboard identifies net salary and comparable rent to COL as factors that are favorable for cities in terms of customers' spending and cost-effectiveness in opening new locations.

An interesting finding from Dashboard 1 is that population has a negative correlation with COL, which means that cities with more customers tend to have a lower cost of living. This implies that big population cities could be a key metric for defining new markets as they could provide a bigger market size and less operating costs for DD.

Dashboard 2 focuses on suggestions from the Business Intelligence team based on data analysis and interpretation. It highlights that the most favorable cities for DD are mostly located in the Asia-Pacific region, including Taiwan, China, South Korea, and Australia. The score systems also target countries with big populations such as China and developed countries in Asia like Japan and South Korea. These findings align with some predictions made by the BI team.

Dashboard 2 shows that European and North American countries, despite having similar properties to DD's current market, may not be as suitable for expansion as predicted based on the analysis of the dashboard.

4.3 Limitations and Weakness

In any project, it's always important to acknowledge its limitations for comprehensive and accurate results and feedback. One of the limitations that this project faced is that the analysis is based solely on secondary data sources, which may have limitations in terms of accuracy and completeness. Another limitation is that our study only considers a limited number of factors that may impact restaurant expansion, and other factors such as cultural and social preferences were not considered, which can greatly impact the success of restaurant chains in different regions. Therefore, future studies should aim to collect more comprehensive and region-specific data to improve the accuracy of the analysis. In addition, the analysis focuses on the current market conditions and does not consider potential changes or disruptions that may occur in the near or far future. Lastly, the study only analyses major cities worldwide, and the findings may not apply to smaller cities or rural areas around the globe.

5. Next Steps

Our analysis has provided some initial insights on the market expansion strategy for DashDelicious using a wide range of data from three datasets, including population, cost of living, meal prices (household and restaurant), rental prices, disposable incomes and purchasing power by countries and cities. We believe that future research might be able to build upon the findings of this study and explore additional factors that can find market potential. One potential method will be to analyze the impact of cultural differences and the impact of different regulatory environments on the restaurant industry. Future studies can incorporate and explore emerging technologies like artificial intelligence and machine learning to analyze and predict consumer behaviors and preferences in different regions, identify gaps in these potential markets, assess market competitors and collect feedback from existing and potential customers. These would help define opportunities and challenges for restaurant expansions regarding go-to-market strategy, personalized offerings development and marketing and business development.

References

Ciaburro, G. (2018). *Regression analysis with R: Design and develop statistical nodes to identify unique relationships within data at scale* (1st edition). Packt. https://oula.finna.fi/Record/oy_electronic_oy.9917615875006252?sid=2943026784

Creative Commons—CC0 1.0 Universal. (n.d.). Retrieved April 22, 2023, from https://creativecommons.org/publicdomain/zero/1.0/

Dedić, N., & Stanier, C. (2017). Towards Differentiating Business Intelligence, Big Data, Data Analytics and Knowledge Discovery. In F. Piazolo, V. Geist, L. Brehm, & R. Schmidt (Eds.), *Innovations in Enterprise Information Systems Management and Engineering* (pp. 114–122). Springer International Publishing. https://doi.org/10.1007/978-3-319-58801-8_10

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, *39*(11), 27–34. https://doi.org/10.1145/240455.240464

Géron, A. (2019). *Hands-on machine learning with Scikit-Learn and TensorFlow: Concepts, tools, and techniques to build intelligent systems* (Second edition). O’Reilly. https://oula.finna.fi/Record/oy_electronic_oy.9917244393906252?sid=2943235053

Hackett, D. W. (1976). The International Expansion of U.S. Franchise Systems: Status and Strategies. *Journal of International Business Studies*, *7*(1), 65–75. https://doi.org/10.1057/palgrave.jibs.8490883

Hedgebeth, D. (2007). Data‐driven decision making for the enterprise: An overview of business intelligence applications. *VINE*, *37*, 414–420. https://doi.org/10.1108/03055720710838498

Kumar, S. M., & Belwal, M. (2017). Performance dashboard: Cutting-edge business intelligence and data visualization. *2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon)*, 1201–1207. https://doi.org/10.1109/SmartTechCon.2017.8358558

Ridzuan, F., & Wan Zainon, W. M. N. (2019). A Review on Data Cleansing Methods for Big Data. *Procedia Computer Science*, *161*, 731–738. https://doi.org/10.1016/j.procs.2019.11.177

Rivas, R. (2011). Internationalization of Peruvian cuisine: An analysis of internationalization strategies of Peruvian restaurants. *Innovar: Revista de Ciencias Administrativas y Sociales*, *21*, 205–216.

Robinson, G. J., & Lundstrom, W. J. (2003). Market expansion strategy: Development of a conceptual market expansion decision scorecard. *Strategic Change*, *12*(5), 259–272. https://doi.org/10.1002/jsc.642

Sharda, R., Delen, D., Turban, E., Aronson, J. E., Liang, T.-P., & King, D. (2019). *Business intelligence, analytics, and data science: A managerial perspective* (Fourth edition). Pearson. https://oula.finna.fi/Record/oy.9917682881606252?sid=2942329179

Sharmiladevi, J. C. (2017). UNDERSTANDING DUNNING'S OLI PARADIGM. *Indian Journal of Commerce and Management Studies*, *VIII*(3), 47–52. https://doi.org/10.18843/ijcms/v8i3/07

Strengholt, P. (2020). *Data Management at Scale: Best Practices for Enterprise Architecture* (1st edition). O'Reilly Media. https://www.oreilly.com/library/view/data-management-at/9781492054771/

Tableau. (2023). *Level of Detail Expressions*. https://help.tableau.com/current/pro/desktop/en-us/calculations_calculatedfields_lod.htm

Vaughn, C. L. (1974). *Franchising: Its nature, scope, advantages and development*. Lexington Books.

Welch, L. S., Benito, G. R. G., & Petersen, B. (2007). *Foreign operation methods: Theory, analysis, strategy*. Edward Elgar.

Appendix 1 – Dashboards

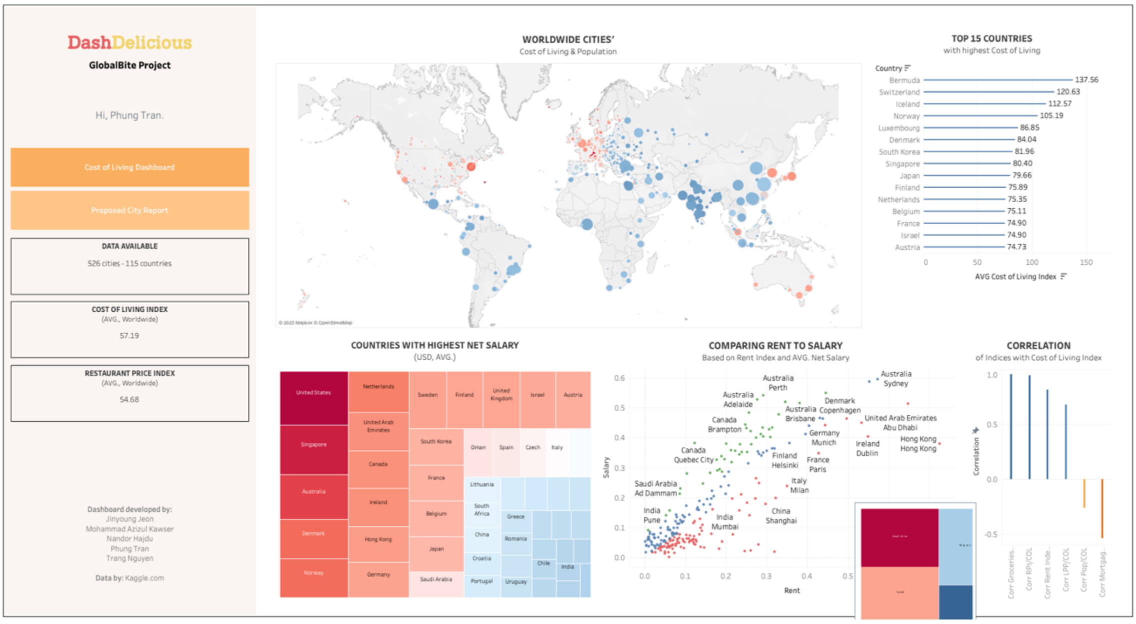

Figure 9. Cost of Living Dashboard

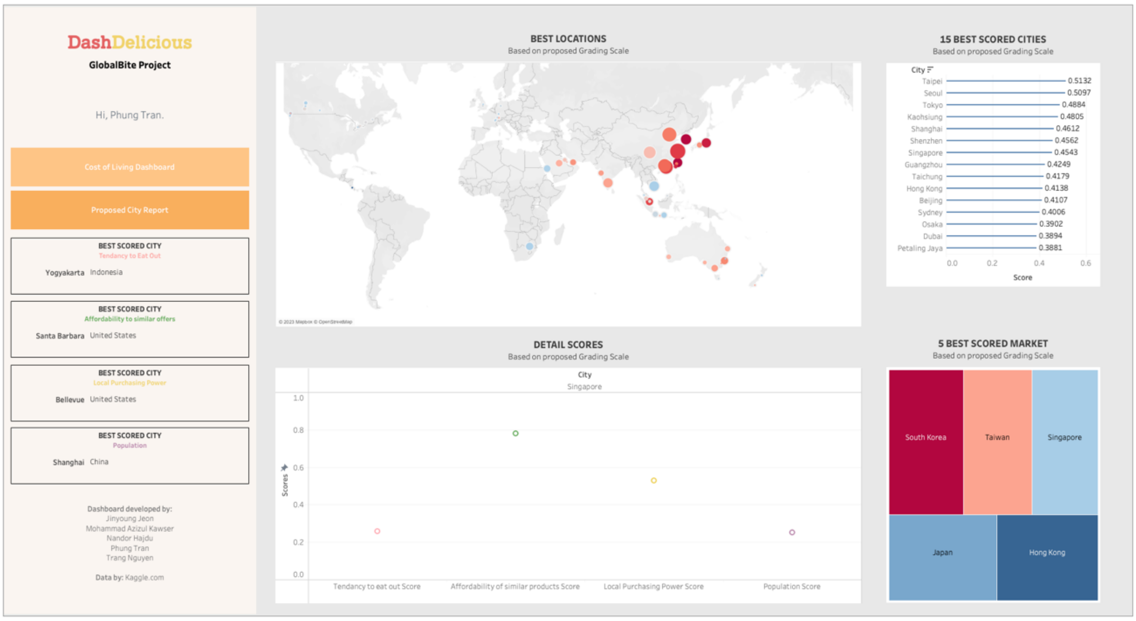

Figure 10. Proposed City Report Dashboard