

Cost Analysis

STUDYING ABROAD



What we'll cover today

1

Problem Statement

2

Statistical Model

3

Literature Review

4

Progress Update

5

Demo

Problem Statement

THE PROBLEM

Prospective international students often struggle to estimate the total cost of studying abroad due to varying expenses like tuition, living costs, and additional fees across different countries and institutions.

THE GOAL

Develop a predictive model to estimate the total cost of studying abroad based on factors such as country, institution, and program duration, aiding students in financial planning.



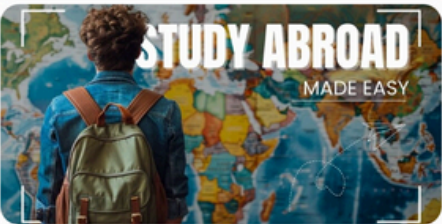
Ridge Regression

Ridge Regression is a regularized version of Multiple Linear Regression. It adds a penalty term to the cost function to discourage large coefficients. In this project, RR will predict the total cost of studying abroad using factors like country, institution, and program duration.

Data set Selection

Cost of International Education

Cost of International Education: Comparative Financial Dataset for Global Study



Data Card Code (24) Discussion (1) Suggestions (0)

About Dataset

This **Cost of International Education** dataset compiles detailed financial information for students pursuing higher education abroad. It covers multiple countries, cities, and universities around the world, capturing the full tuition and living expenses spectrum alongside key ancillary costs. With standardized fields such as tuition in USD, living-cost indices, rent, visa fees, insurance, and up-to-date exchange rates, it enables comparative analysis across programs, degree levels, and geographies. Whether you're a prospective international student mapping out budgets, an educational consultant advising on affordability, or a researcher studying global education economics, this dataset offers a comprehensive foundation for data-driven insights.

Usability ⓘ

10.00

License

[Attribution 4.0 International \(CC ...](#)

Expected update frequency

Annually

Tags

Education

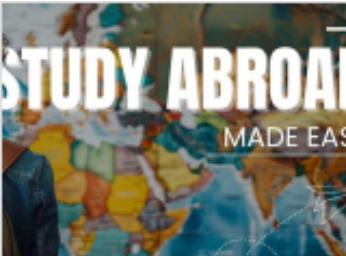
Universities and Colleges

Currencies and Foreign Exchange

Description

Column	Type	Description
Country	string	ISO country name where the university is located (e.g., "Germany", "Australia").
City	string	City in which the institution sits (e.g., "Munich", "Melbourne").

View more



Cost of International Education

Cost of International Education: Comparative Financial Dataset for Global Study



Literature Review

Key Studies on Cost of Foreign Education

1. Wiley Study (2025).
2. IAS Services Report (2025).
3. HEPI Analysis (2023).
4. NACE Journal (2021).
5. EconomiX Working Paper (2017).
6. Taylor & Francis Study (2023).

Key Findings

- Tuition fees inversely impact mobility: Italy saw reduced enrollment with fee hikes, mitigated by education quality.
- US international student enrollment dropped 10% (2015–2019) due to rising costs and limited financial aid.
- Living costs for international students: £900–£1,400/month in the UK, exacerbating financial stress.
- US tuition for international students rose \$5,000 (2012–2016) vs. \$450 for domestic peers.
- Experiential learning (e.g., UK MBA programs) improved employability for 95% of Indian subcontinent students.
- Host capacity, education ROI, and local costs are key determinants of student destination choices.

Exploratory Data Analysis (EDA)

2. EXPLORATORY DATA ANALYSIS

=====

Dataset Info:
<class 'pandas.core.frame.DataFrame'>
Index: 922 entries, 0 to 999
Data columns (total 10 columns):

#	Column	Non-Null Count	Dtype
0	Country	922 non-null	object
1	University_Ranking	922 non-null	int64
2	Tuition_Fee_USD	922 non-null	float64
3	Living_Cost_USD	922 non-null	float64
4	GDP_Per_Capita	922 non-null	float64
5	University_Type	922 non-null	object
6	Program_Duration_Years	922 non-null	int64
7	City_Tier	922 non-null	int64
8	Language_Requirement_Score	922 non-null	float64
9	Total_Annual_Cost_USD	922 non-null	float64

dtypes: float64(5), int64(3), object(2)
memory usage: 79.2+ KB
None

Descriptive Statistics:

	University_Ranking	Tuition_Fee_USD	Living_Cost_USD	GDP_Per_Capita \
count	922.000000	922.000000	922.000000	922.000000
mean	249.288503	26554.393989	16392.279377	44728.416371
std	144.798843	14145.682995	7467.040758	19896.412911
min	1.000000	111.330724	153.297832	-19788.754499
25%	128.000000	16291.144870	11228.024648	31535.757665
50%	254.000000	25970.876103	16231.032410	44210.648222
75%	369.750000	36703.162768	21044.934634	58589.970977
max	499.000000	71410.735234	41083.928646	102684.645531

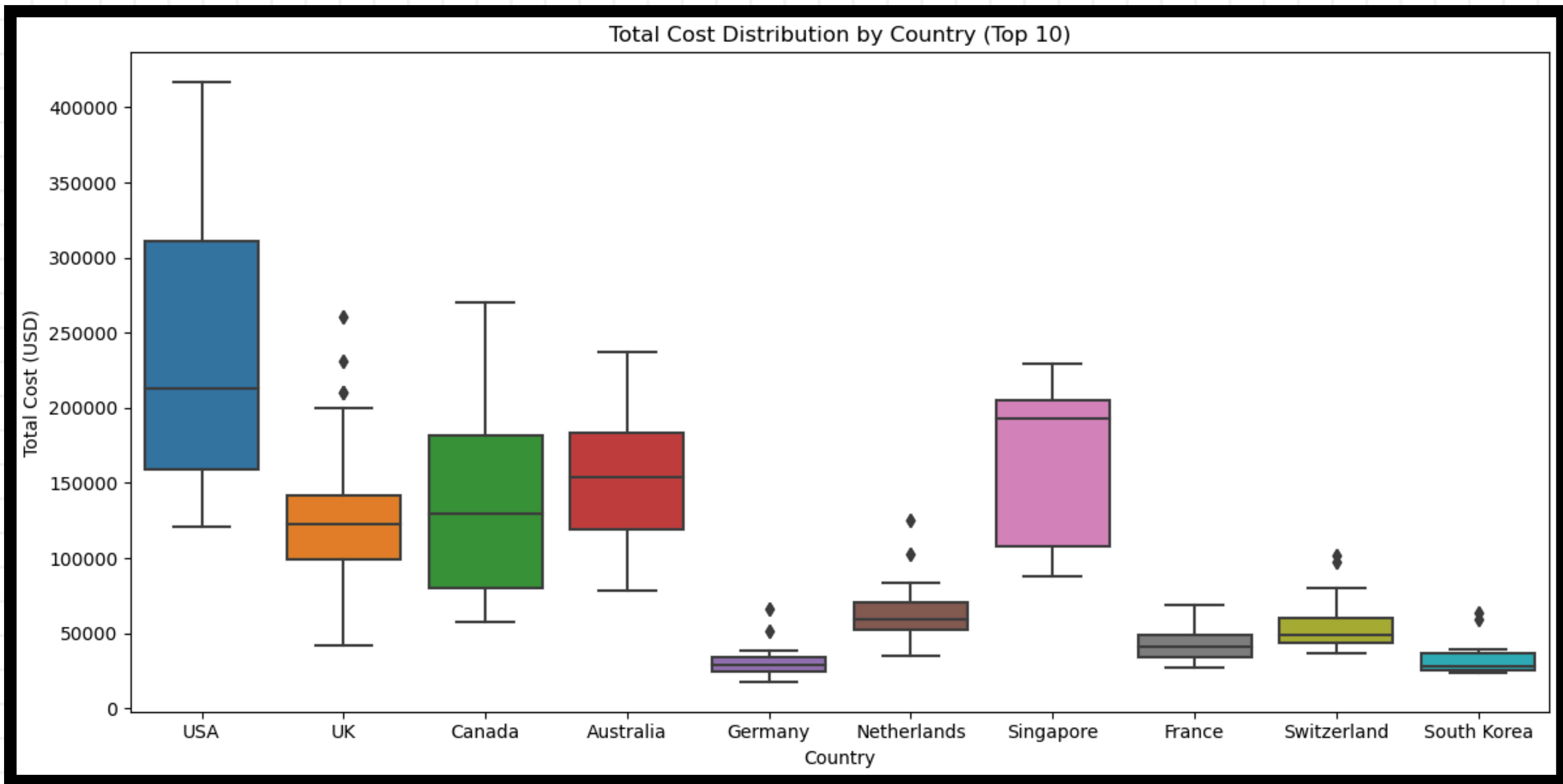
	Program_Duration_Years	City_Tier	Language_Requirement_Score \
count	922.000000	922.000000	922.000000
mean	2.425163	2.014100	6.543715
std	1.088553	0.813043	1.528399
min	1.000000	1.000000	1.555620
25%	1.000000	1.000000	5.498246
50%	2.000000	2.000000	6.574735
75%	3.000000	3.000000	7.596381
max	4.000000	3.000000	10.698469

	Total_Annual_Cost_USD
count	922.000000
mean	42971.353254
std	16189.820737
min	3198.229733
25%	31299.606715
50%	42431.038663
75%	53346.254345
max	97485.058625

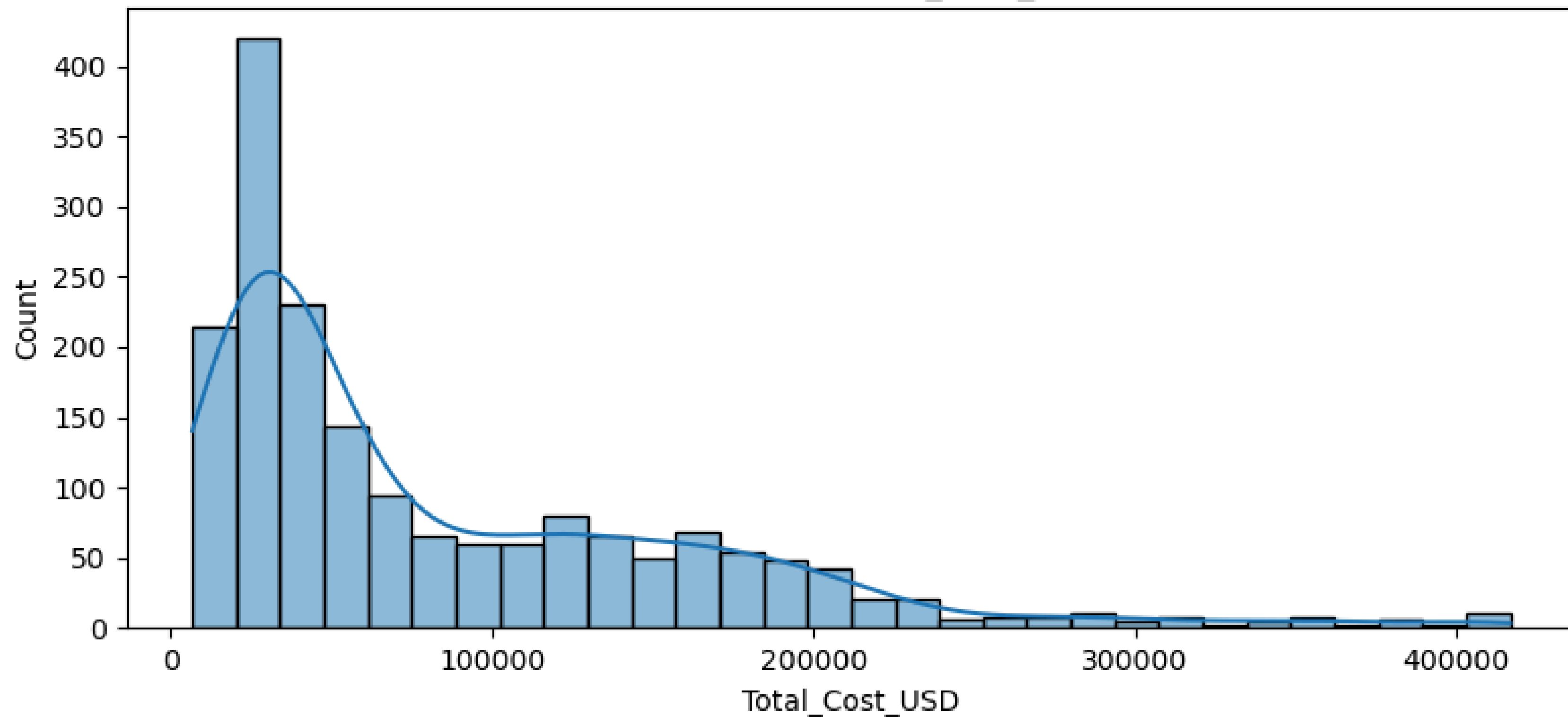
Missing Values:

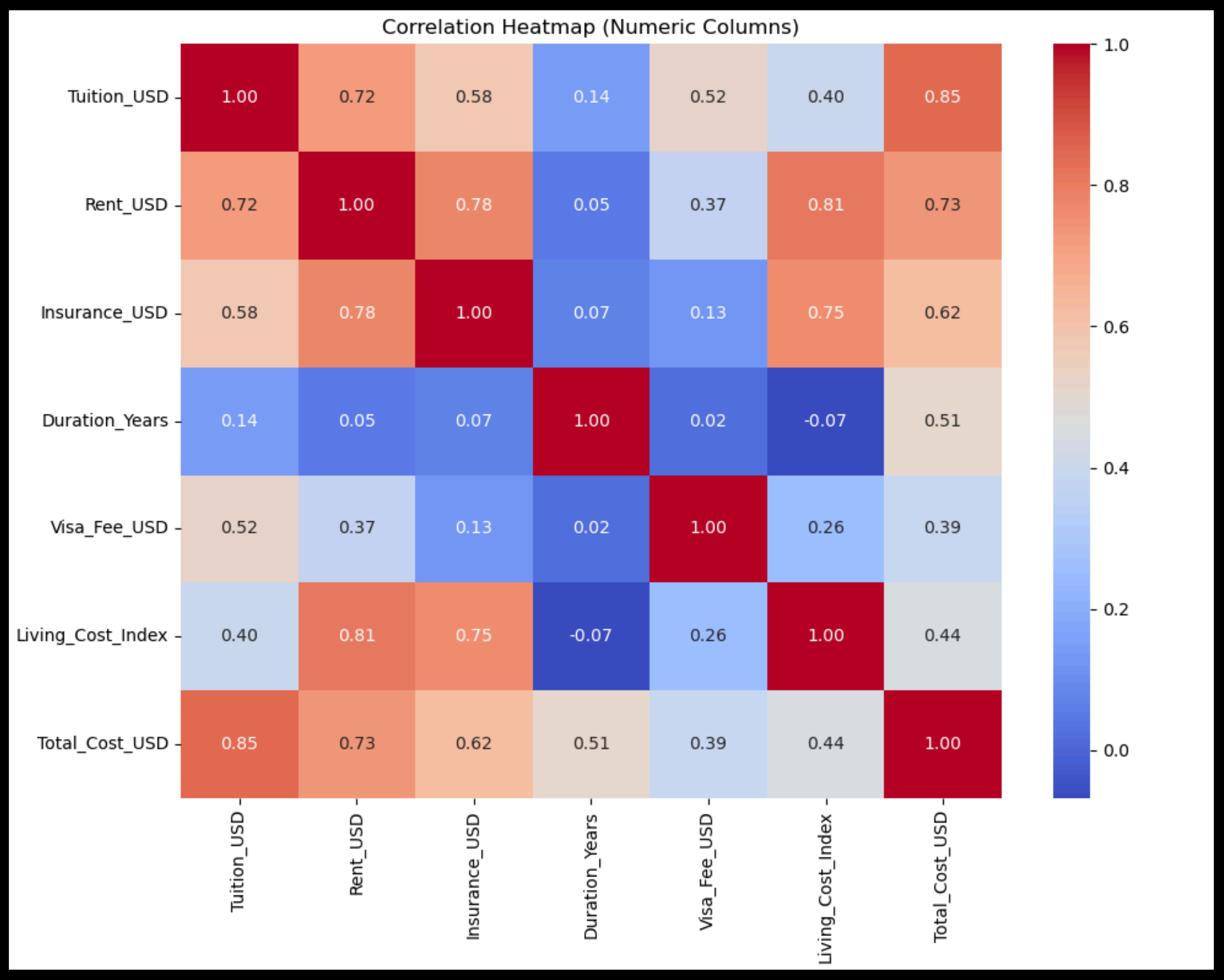
Country	0
University_Ranking	0
Tuition_Fee_USD	0
Living_Cost_USD	0
GDP_Per_Capita	0
University_Type	0
Program_Duration_Years	0
City_Tier	0
Language_Requirement_Score	0
Total_Annual_Cost_USD	0

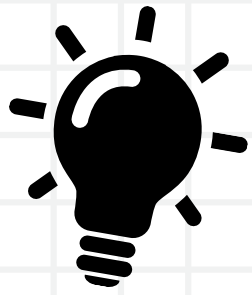
dtype: int64



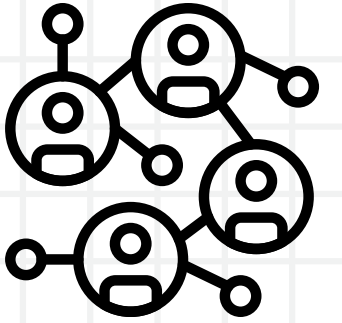
Distribution of Total_Cost_USD







Statistical Analysis & Validation

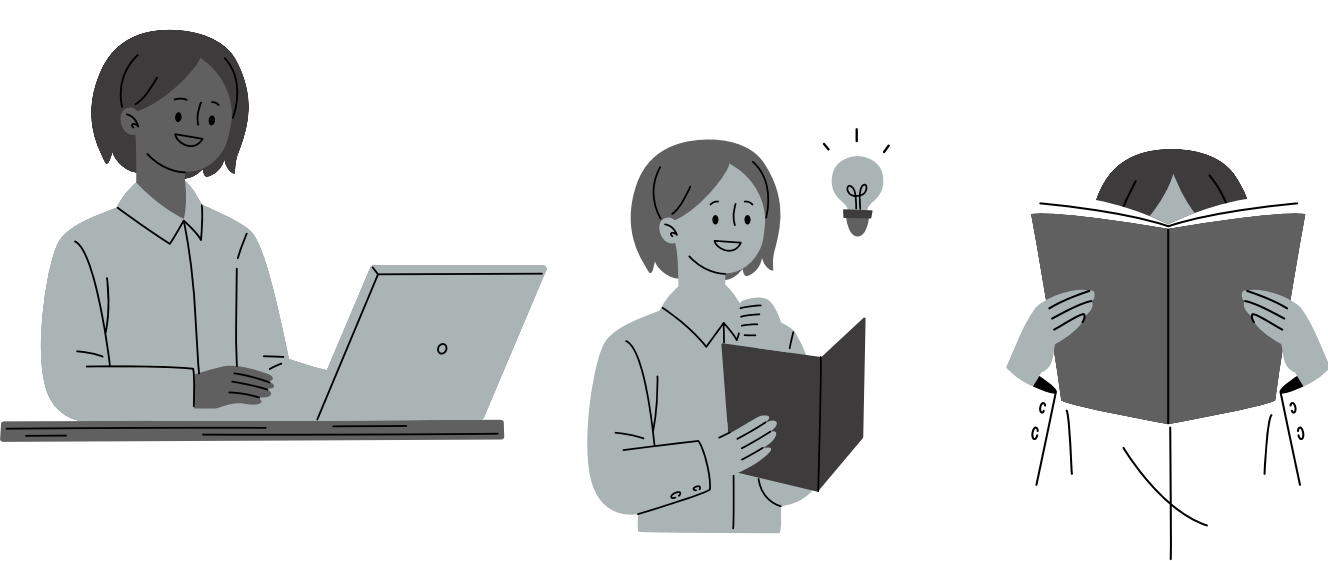


```
--- Key Findings ---
Countries in dataset: 73
Programs: 93
Degree levels: ['Master' 'Bachelor' 'PhD' nan 'Level']
Average total cost (USD): 85,098
Median total cost (USD): 51,008
Country with most entries: UK

--- OLS Regression Summary (MLR) ---
                        OLS Regression Results
=====
Dep. Variable:          Total_Cost_USD    R-squared:                0.917
Model:                  OLS              Adj. R-squared:         0.917
Method:                 Least Squares    F-statistic:             3320.
Date:                   Tue, 24 Jun 2025  Prob (F-statistic):      0.00
Time:                   08:14:50         Log-Likelihood:          -20754.
No. Observations:      1814             AIC:                    4.152e+04
Df Residuals:          1807             BIC:                    4.156e+04
Df Model:               6
Covariance Type:        nonrobust
=====
                        coef      std err          t      P>|t|      [0.025      0.975]
-----
const                -9.724e+04    3902.151    -24.920    0.000    -1.05e+05    -8.96e+04
Tuition_USD           2.7645         0.063     44.049    0.000         2.641         2.888
Rent_USD              37.6198         2.681     14.033    0.000         32.362         42.878
Insurance_USD         18.5190         3.167      5.848    0.000         12.308         24.730
Duration_Years       3.342e+04    575.855     58.027    0.000     3.23e+04     3.45e+04
Visa_Fee_USD        -13.1626         4.785     -2.751    0.006        -22.547         -3.779
Living_Cost_Index    -81.8893        84.856     -0.965    0.335        -248.315         84.536
=====
Omnibus:              62.325    Durbin-Watson:           2.193
Prob(Omnibus):        0.000    Jarque-Bera (JB):       165.704
Skew:                 0.064    Prob(JB):               1.04e-36
Kurtosis:             4.475    Cond. No.:              1.74e+05
=====
```

Application Demo





Thank You!

