

1. INTRODUCTION

Aim and Scope:

The objective of this project is to apply the knowledge acquired within the field of Geographic Information Systems. In this context, relevant data from multiple sources will be collected for the ten cities assigned in the project, and a coherent narrative will be developed based on these data. Subsequently, all procedures specified in the assignment guidelines will be carried out using the QGIS software, including query operations and spatial analyses. This project aims to analyze the tourism potential and spatial distribution of tourism-related indicators in selected provinces of Türkiye, namely **Erzurum, Artvin, Tokat, Sivas, Samsun, Rize, Trabzon, Ordu, Erzincan, and Bayburt.**

The Story:

The study focuses on evaluating the relationship between the number of places to visit, domestic and foreign tourist arrivals, and accommodation capacity across these provinces. Within the scope of the study, a tourism potential coefficient and a hostel coefficient were derived using tourism-related attributes in order to better understand the relative tourism performance of each province. These derived indicators enable a comparative spatial analysis by normalizing tourism demand with respect to accommodation infrastructure and touristic attractions. Understanding the spatial variation of tourism potential and the influence of foreign and domestic tourists is crucial for regional tourism planning, sustainable development strategies, and investment prioritization. By integrating statistical tourism data with geographic information systems (GIS), this project provides a comprehensive spatial perspective on tourism dynamics in the selected provinces.

Through GIS-based analyses, tourism-related data were assessed spatially, provinces were compared using derived indicators, and regions that may be considered priorities for future investment were identified. Consequently, this project demonstrates the importance of spatial analysis in tourism planning and decision-making processes, offering a data-driven perspective for sustainable tourism development.

Tools and Software:

In this project, QGIS was utilized as the primary software for conducting georeferencing, digitization, and spatial analysis processes. The software enabled the efficient management, visualization, and analysis of spatial data in accordance with the project objectives.

2. DATA COLLECTION AND PREPARATION

Data Sources:

The data utilized in this project were obtained from the following reliable and official sources:

1. ***TURKISH STATISTICAL INSTITUTE (TÜİK):***

Population statistics and tourism-related data, including domestic and foreign tourist numbers.

2. ***MINISTRY OF CULTURE AND TOURISM (KTB):***

Data on accommodation facilities, including the number of hotels and tourism infrastructure.

3. ***MINISTRY OF CULTURE AND TOURISM – Cultural Portal (Kültür Portalı):***

Information on touristic and cultural attractions, including the number of places to visit in each province.

These data sources were selected due to their comprehensive coverage, official status, and reliability. The integration of statistical tourism data with cultural and accommodation information enables a comprehensive spatial analysis of tourism potential and regional differences across the selected provinces.

Georeferencing:

Georeferencing is the process of assigning real-world geographic coordinates to raster data to accurately position them within a geographic coordinate system. In this project, raster maps of the selected provinces were georeferenced in QGIS to ensure consistency with vector datasets and attribute information. Clearly identifiable control points, such as provincial boundaries and intersection points visible on both raster and vector layers, were used to align the raster data with real-world coordinates. An appropriate coordinate reference system (CRS) was selected to maintain compatibility with other datasets, and the accuracy of the georeferencing was verified by overlaying the raster with vector boundary layers. This process enabled accurate digitization and reliable spatial analysis.

Digitization:

Digitization is the process of converting georeferenced raster data into vector format for spatial analysis in a GIS environment. In this study, provincial boundaries of the selected study area were manually digitized as polygon layers using QGIS. The digitized layers were then enriched with tourism-related attributes, including the number of places to visit, domestic and foreign tourist numbers, and accommodation facilities. These vector layers were subsequently used for spatial analyses, queries, and map visualizations.

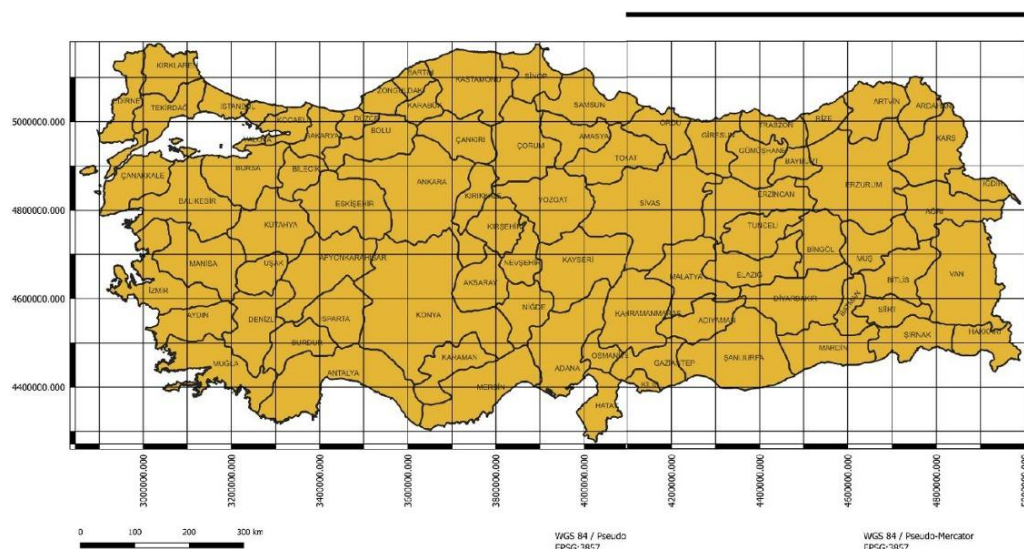


Figure 1 : Digitized vector layer representing provincial boundaries



Figure 2 : Location of the selected provinces in Türkiye

3. DATABASE DESIGN AND ATTRIBUTES

Database Design:

During the data collection and processing phase of the project, both spatial data and descriptive attribute data were systematically compiled for the ten provinces included in the study area. As stated in the group project proposal, raw data obtained from sources such as the Turkish Statistical Institute (TÜİK), the Ministry of Culture and Tourism, and the Culture Portal were integrated with digitized vector data to establish a comprehensive database within the GIS environment. The collected attributes specifically included the number of hotels, the number of touristic attractions to visit, domestic tourist numbers, and foreign tourist numbers for each province. By associating these datasets with their corresponding spatial features, a structured and consistent geodatabase was created, enabling effective querying, analysis, and interpretation of the tourism potential across the selected provinces.

Attribute Table:

Attribute Table serves as the main data structure in which numerical and descriptive information collected for each province is stored in an organized manner. Within the scope of this project, data related to population, tourism, accommodation facilities, and other relevant indicators were entered into the attribute table for each province. This table enabled the comparison of provinces based on variables such as domestic and foreign tourist numbers, hotel counts, and derived

coefficients. Furthermore, the attribute table formed the basis for attribute queries, spatial analyses, and the interpretation of the results.

	fid	GID_1	COUNTRY	NAME_1	TO_VISIT	DOMESTIC_TOURIST	FOREIGN_TOURIST	HOTELS	POTENTIAL_COEFFICIENT	HOSTEL_COEFFICIENT	
1		NULL	TUR.31_1	Turkey	Erzurum	59	637757	47949	98	20935.57475	213.6283137755102
2		1	TUR.10_1	Turkey	Artvin	78	339587	28530	126	14912.898000000001	118.35633333333334
3		1	TUR.74_1	Turkey	Tokat	40	344699	5615	65	7062.43	108.65276923076924
4		1	TUR.72_1	Turkey	Sivas	31	253457	7976	72	4114.0255	57.13924305555555
5		1	TUR.67_1	Turkey	Samsun	38	352089	39329	145	7810.5675	53.86598275862069
6		1	TUR.65_1	Turkey	Rize	63	227169	18035	178	8007.977250000002	44.98863623595506
7		1	TUR.75_1	Turkey	Trabzon	71	271079	145513	440	17371.87175	39.48152670454545
8		1	TUR.63_1	Turkey	Ordu	83	145424	1883	166	6152.31275	37.062125
9		1	TUR.30_1	Turkey	Erzincan	10	215300	10076	53	1152.07	21.737169811320754
10		0	TUR.15_1	Turkey	Bayburt	13	12315	140	8	81.41250000000001	10.176562500000001

Figure 3 : Attribute table of tourism indicators used in the analysis

Derived Attributes:

In addition to the original variables collected from various data sources, several derived attributes were calculated and added to the attribute table in order to enhance the analytical capability of the dataset. These derived datas were created by combining existing tourism-related variables to allow a more meaningful comparison between provinces. The Tourism Potential Coefficient was calculated by relating the total number of domestic and foreign tourists to the number of places to visit in each province. This coefficient provides an indicator of tourism intensity and demand relative to touristic attractions. Furthermore, the Hostel Coefficient was derived by normalizing tourism demand with respect to accommodation capacity, using the number of hotels as a reference. This coefficient enables the evaluation of whether the existing accommodation infrastructure is sufficient relative to tourism demand. The inclusion of these derived attributes allowed for comparative spatial analyses, supported attribute queries, and contributed to the interpretation of tourism performance across the selected provinces.

$$\text{POTENTIAL_COEFFICIENT} = 0.1 \times \text{TO_VISIT} \times (0.005 \times \text{DOMESTIC_TOURIST} + 0.0075 \times \text{FOREIGN_TOURIST})$$

$$\text{HOSTEL_COEFFICIENT} = \frac{\text{POTENTIAL_COEFFICIENT}}{\text{HOTELS}}$$

Ezgi Yaman's First Query

Query: Which provinces have high tourism potential but insufficient accommodation infrastructure?

Explanation:

The query aims to identify specific provinces that meet two distinct criteria: high demand (visitor volume) and a potential shortage or specific characteristic in accommodation infrastructure (low hostel/budget density or a specific coefficient threshold).

```
( coalesce("DOMESTIC_TOURIST",0) +  
coalesce("FOREIGN_TOURIST",0) ) >  
250000  
AND  
coalesce("HOSTEL_COEFFICIENT",0) >  
100
```

Figure 12: Query Conditions

The Logic Breakdown:

- **High Tourism Volume:** The expression `coalesce("DOMESTIC_TOURIST", 0) + coalesce("FOREIGN_TOURIST", 0) > 250,000` calculates the "Total Tourist Volume". By using `coalesce`, the query ensures that even if one value (Domestic or Foreign) is missing (NULL), it is treated as zero rather than breaking the calculation. It filters for provinces with more than 250,000 total visitors.
- **Infrastructure Sensitivity:** The condition `coalesce("HOSTEL_COEFFICIENT", 0) > 100` targets provinces where the Hostel Coefficient exceeds a specific threshold. In this context, a high coefficient suggests that the demand for budget-friendly or alternative accommodation significantly outweighs the current supply, or that the province's capacity is heavily skewed.

Results:

The resulting list highlights provinces ripe for investment. These are areas where tourism is booming, but the accommodation infrastructure has not yet scaled to meet the specific needs of the visiting demographic.

	fid	GID_1	COUNTRY	NAME_1	TO_VISIT	DOMESTIC_TOURIST	FOREIGN_TOURIST	HOTELS	POTENTIAL_COEFFICIENT	HOSTEL_COEFFICIENT
1	31	TUR.31_1	Turkey	Erzurum	59	637757	47949	98	20935.57475	213.6283137755102
2	10	TUR.10_1	Turkey	Artvin	78	339587	28530	126	14912.898000000001	118.35633333333334
3	74	TUR.74_1	Turkey	Tokat	40	344699	5615	65	7062.43	108.65276923076924
4	72	TUR.72_1	Turkey	Sivas	31	253457	7976	72	4114.0255	57.13924305555555
5	67	TUR.67_1	Turkey	Samsun	38	352089	39329	145	7810.5675	53.86598275862069
6	65	TUR.65_1	Turkey	Rize	63	227169	18035	178	8007.977250000002	44.98863623595506
7	75	TUR.75_1	Turkey	Trabzon	71	271079	145513	440	17371.87175	39.48152670454545
8	63	TUR.63_1	Turkey	Ordu	83	145424	1883	166	6152.31275	37.062125
9	30	TUR.30_1	Turkey	Erzincan	10	215300	10076	53	1152.07	21.737169811320754
10	15	TUR.15_1	Turkey	Bayburt	13	12315	140	8	81.41250000000001	10.176562500000001

Figure 13: Comparison of Tourism Demand and Accommodation Capacity by Province

Ezgi Yaman's Second Query

Query: Which province represents the absolute maximum infrastructure bottleneck relative to tourist demand?

Explanation:

This query performs a relative capacity analysis to find the province where the gap between tourist demand and accommodation supply is at its absolute maximum.

```
(
  (coalesce("DOMESTIC_TOURIST",0) +
  coalesce("FOREIGN_TOURIST",0))
  / "HOTELS"
)
=
maximum(
  (coalesce("DOMESTIC_TOURIST",0) +
  coalesce("FOREIGN_TOURIST",0))
  / "HOTELS"
)
```

Figure 14: Query Conditions

The Logic Breakdown:

- **Total Tourist Demand:** It first sums the domestic and foreign visitors `coalesce("DOMESTIC_TOURIST", 0) + coalesce("FOREIGN_TOURIST", 0)`, ensuring that missing data (NULLs) are treated as zero.
- **Load Ratio (Tourists per Hotel):** By dividing the total volume by the number of hotels ("`HOTELS`"), the query calculates the "Tourists per Hotel" ratio. This metric represents the density of pressure on existing infrastructure.
- **Maximum Filter:** The `maximum(...)` function acts as a global filter. Instead of showing all provinces, it isolates the record(s) where this ratio is the highest in the entire dataset.

Results:

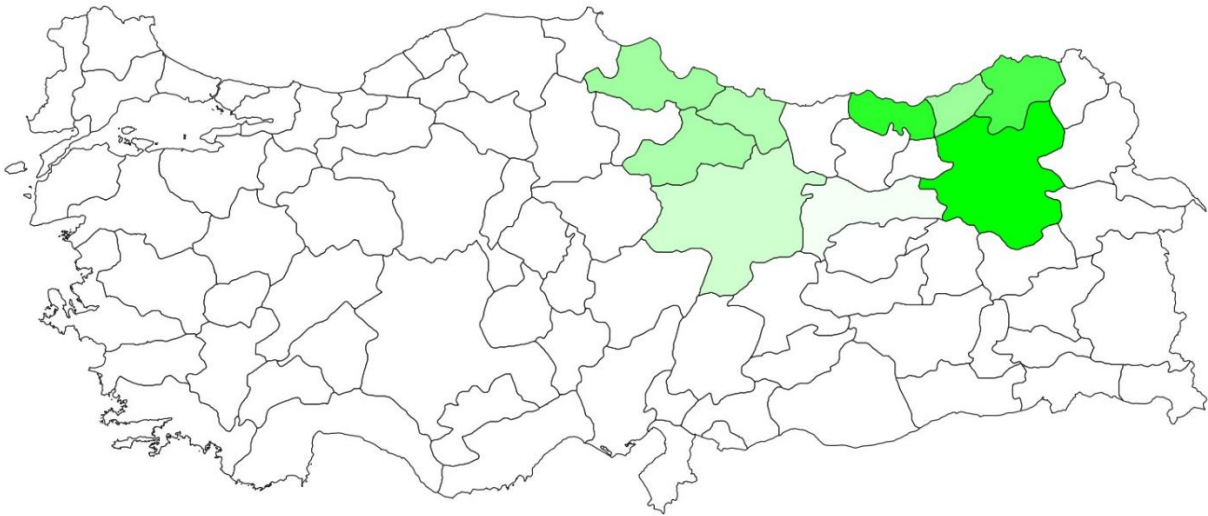
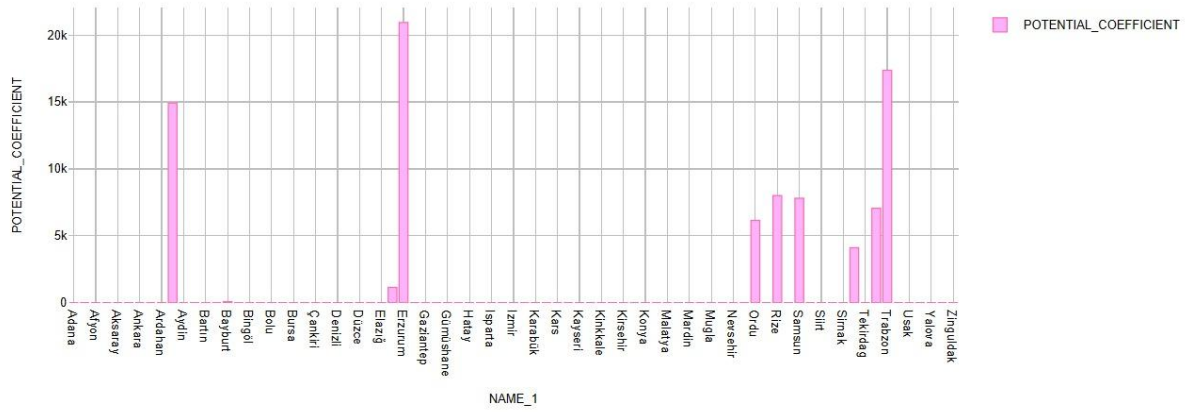
Based on the data provided in the table, the query would return **Erzurum** as the result. While provinces like **Trabzon** have a high visitor volume (~416,592), they also have a much higher hotel count (440), resulting in a much lower load ratio (~946 tourists per hotel). **Tokat** and **Artvin** also show significant stress, but their ratios (~5,389 and ~2,921 respectively) do not reach the peak seen in Erzurum.

Erzurum represents the absolute maximum pressure point in the dataset. From a strategic standpoint, this province has the highest "Infrastructure Deficit," making it the most lucrative and necessary location for further hospitality development.

	fid	GID_1	COUNTRY	NAME_1	TO_VISIT	DOMESTIC_TOURIST	FOREIGN_TOURIST	HOTELS	POTENTIAL_COEFFICIENT	HOSTEL_COEFFICIENT
1	31	TUR.31_1	Turkey	Erzurum	59	637757	47949	98	20935.57475	213.6283137755102
2	10	TUR.10_1	Turkey	Artvin	78	339587	28530	126	14912.898000000001	118.35633333333334
3	74	TUR.74_1	Turkey	Tokat	40	344699	5615	65	7062.43	108.65276923076924
4	72	TUR.72_1	Turkey	Sivas	31	253457	7976	72	4114.0255	57.13924305555555
5	67	TUR.67_1	Turkey	Samsun	38	352089	39329	145	7810.5675	53.86598275862069
6	65	TUR.65_1	Turkey	Rize	63	227169	18035	178	8007.977250000002	44.98863623595506
7	75	TUR.75_1	Turkey	Trabzon	71	271079	145513	440	17371.87175	39.48152670454545
8	63	TUR.63_1	Turkey	Ordu	83	145424	1883	166	6152.31275	37.062125
9	30	TUR.30_1	Turkey	Erzincan	10	215300	10076	53	1152.07	21.737169811320754
10	15	TUR.15_1	Turkey	Bayburt	13	12315	140	8	81.41250000000001	10.176562500000001

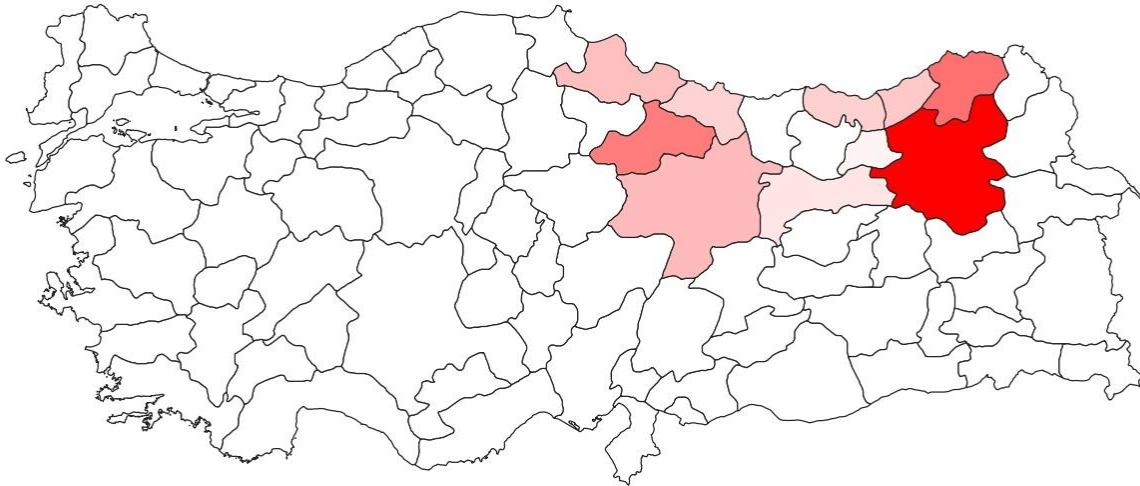
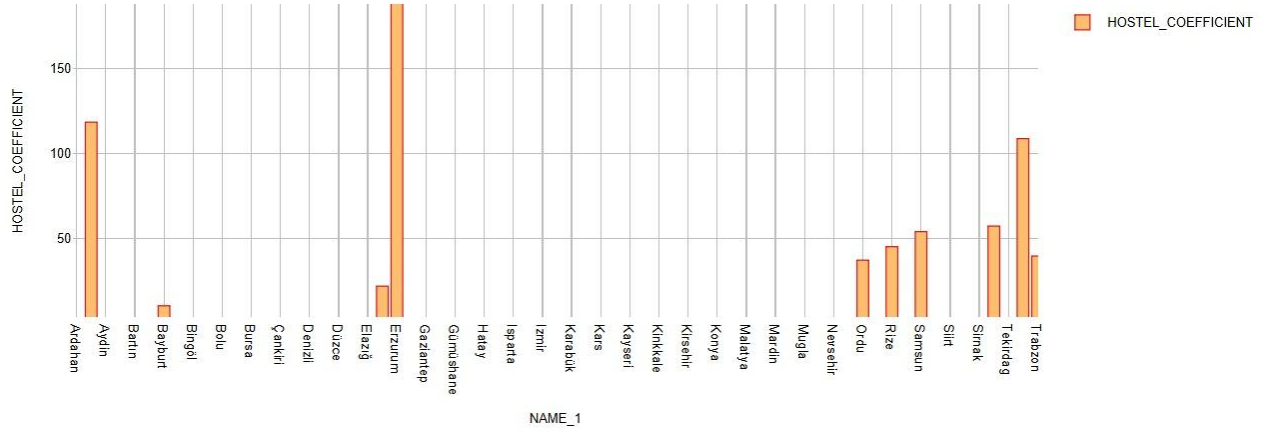
Figure 15: Province with the Highest Ratio of Tourism Demand to Existing Accommodation Infrastructure

5. Graphical Analysis



Figures 16 & 17: Tourism Potential Analysis

According to the analysis, the provinces with the highest tourism potential are Erzurum, Artvin and Trabzon, respectively.



Figures 18 & 19: Hotel Investment Analysis

The analysis highlights Erzurum, Artvin and Trabzon are also as the most promising cities for hotel investment.

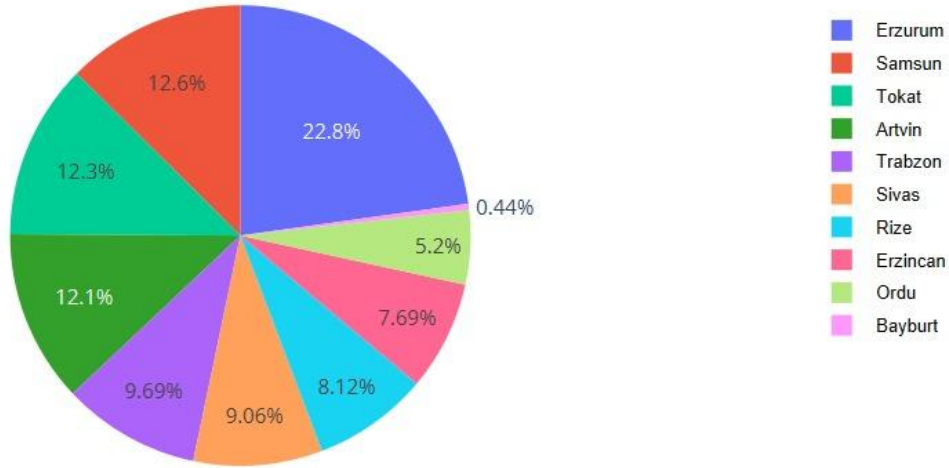


Figure 20: Domestic Tourist Rates

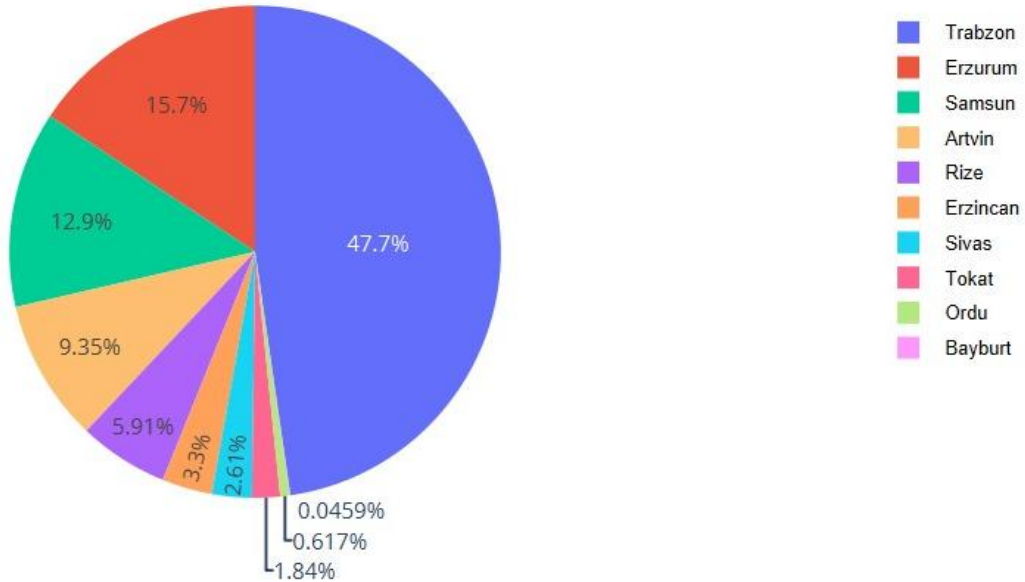


Figure 21: Foreigner Tourist Rates