**A blue sign with white text

AI-generated content may be incorrect.**

**DSA 210**

**RELATIONSHIP BETWEEN CULTURAL ACTIVITIES AND SUICIDE RATE IN TURKIYE**

**PROJECT REPORT**

Eren Emir Ayar

32206

Date: 30.05.2025

Instructor

Selim Balcısoy

**INTRODUCTION**

The purpose of this project is to shed light on possible benefits of cultural activities such as theatre, opera and concerts on the mental health of the people. I chose this topic because understanding the relationship between them may lead to new treatment methods for suicidal people. In addition to that artists in modern Türkiye may be more respected than they are now. I also have personal interest in this topic since I am also into theatre and art.

Search Question: "Do depression rate in Türkiye and attendance rate to any art activity in Türkiye are correlated?"

**WHAT DID I DO?**

After I chose the topic, I looked for data that may be beneficial to use. I conduct detailed research in WHO, Türkiye Ministry of Health and TÜİK databases. I found usable data in TÜİK.

When I found the data, they were filled with irrelevant data. For example, number of suicides in each city in Türkiye, opera attendance in different regions, population in terms of age and gender etc. I used cinema attendance data, theatre attendance data, opera attendance data, population data and suicide data. Each data was between the years 2000-2023. So, I collected them together and purified them from irrelevant information. There was no population data between the years 2001-2006, so I filled these years with linear interpolation.

After I fill population, I calculated suicide rate for each year. Then, I created a table to see the data altogether, and I created a table for suicide rate, theatre, cinema and opera.

When the visualization was over, I made hypothesis testing via p-values and Pearson coefficients. However, I needed to purify the data again due to COVID effect on suicide and cultural activities.

Lastly, I used Naïve Bayes method for Machine Learning part since it was the best option among all ML methods.

**DATA AND GRAPHS**

1. **Data**

**A table with numbers and letters

AI-generated content may be incorrect.**

The data shows us the population, suicide number, theatre attendance, opera attendance, cinema attendance and calculated suicide rate data for the years 2000-2023. Population data for 2001-2006 are estimated with linear interpolation.

1. **Graphs For All Data**

**A group of colorful bars

AI-generated content may be incorrect.**

These graphs show us all the data visually. The main point of these graphs is to indicate the significant drop in attendances and slight rise in suicide rate in the years 2020-2021. This is the key reason why I neglected those 2 years due to COVID.

1. **Scatter Plot with Trend Lines**

A graph with green dots

AI-generated content may be incorrect.

When we look at the scatter plots, we see a slight positive correlation between suicide rates and theater (also cinema). However, opera does not show any significant correlation. It must be analyzed better to reach healthier results.

1. **ROC Curve for Naïve Bayes**

A graph with a line

AI-generated content may be incorrect.

Even though all Naïve Bayes metrics became 1.00, we can’t conclude that it will perfectly work. Since dataset is small, it may be overfitting, and we can’t trust the machine will predict correctly.

**HYPOTHESIS TESTING**

Null Hypothesis (H0): There is no correlation between cultural activities and suicides.

Alternative Hypothesis (H1): There is a negative correlation between cultural activities and suicides.

Since I believed that there is a positive correlation, I built my hypothesis this way. Then, I checked for p-values and Pearson Coefficients.

|  |
| --- |
| Pearson Correlation Coefficient (Theatre): 0.3997 |
| p-value (Theatre): 0.0653 |
| Pearson Correlation Coefficient (Opera): 0.1165 |
| p-value (Opera): 0.6057 |
| Pearson Correlation Coefficient (Cinema): 0.3393 |
| p-value (Cinema): 0.1224 |

Since all p-values are greater than 0.05, we fail to reject the null hypothesis.

According to Pearson Correlation Coefficients, there is only a slight positive relation between each activity and suicide rate. When we consider all, weak correlations may be due to other factors or by chance since coefficients are too close to 0.

**CONCLUSION**

When we look at all the data, graphs and hypothesis testing we can’t see a reliable negative correlation between suicide rates and cultural activities. This may be due to shortcomings of dataset or they may not be a correlation at all. The null hypothesis being failed to reject and seeing no strong correlation indicates that there is no correlation between suicide rates and cultural activities, but it is not certain.

**FUTURE WORK AND SHORTCOMINGS OF THIS PROJECT**

Since we can’t see a strong correlation, we can conclude that there may be some shortcomings of this project. For example, including only 3 cultural activities is not sufficient. If there were more activities, we may have reached a better conclusion. In addition to that, using only the years 2000-2023 may be insufficient. I filled the population of 6 years with linear interpolation and neglected 2 years data due to COVID. It would be better if I could reach more years, more activities. Then I could train the machine better and find more reliable results.

Future work could focus on expanding the dataset with more years and additional features such as other cultural activities. It should also be acknowledged that suicides happen due to too many other factors like social, economic, natural, psychological etc. This project aimed to understand the effects of cultural activities on suicide rates. Using broader data, better ML methods and maybe analyzing the worldwide data will create a future for this work.