Introduction

The economy of a country can be influenced by different sectors such as manufacturing, agriculture, construction, tourism, among others.

One of the countries which has been positively affected by tourism is Iceland. It is a Nordic country located in the Atlantic Ocean. The variety of its landscapes, like geysers, volcanoes, hot springs, and glaciers, catch the attention from people all over the world, regardless of which season it is.

In the last decades, Iceland's economy has been diversifying into manufacturing and service industries, where tourism belongs to. Although it was hugely negatively affected by the COVID-19 pandemic, the amount of tourists increased even more after it.

As we can see in the following graphic that shows the “Percentage change in the number of international visitor arrivals to Iceland from 2010 to 2022”, the international tourism volume has experienced every year growth of 146% in 2022.

Tourism in Iceland was not a big topic to talk about 10 years ago. Due to the develop of this area has grown up too fast, it can be important to predict how many tourists will visit the island in the upcoming years, so that the country, the airport, accommodation, mean of transport, travel agencies, among others, can get prepared in a better way to receive the new visitors and they can also get a good experience on their trip.

A graph with blue and white bars

Description automatically generated

<https://www.statista.com/statistics/694121/tourism-industry-international-visitor-growth-iceland/#:~:text=The%20international%20tourism%20volume%20in,(COVID%2D19)%20pandemic>.

Objectives

The main purpose of the following assignment is to predict how many tourists will visit Iceland per year in the next following 5 years. Each year will be split into two main seasons (Winter- Autumn/ Summer- Spring).

The hypothesis is that every year the amount of visitors will increase exponentially. As a consequence of this increase, the country needs to be prepared in order to welcome the tourists. The knowledge of the number of visitors will have a positive impact on different facilities, because they can get prepared in a better way to receive them.

Also, this can help to optimize the budget that each facility can spend each year to improve their service, as well as to know the amount of employees they need to hire.

On the other hand, a better marketing strategy can be implemented if companies know the number of tourists it is expected to be.

Problem definition

Due to the tourism in Iceland is a new topic that has been more mentioned in the last ten years and it has increased exponentially increase the last years it may be a problem for the country in different aspects such as, budget, infrastructure, services among others.

If the amount of visitors continues increasing exponentially, the country needs to get more knowledge on how to approach this new area.

It is important to address this problem because it can avoid a saturation of sources in the facilities. For instance, the international airport could have a better flow on the landings and arrivals so that there will not be delays. Accommodation, such as hotel, hostel, B&B, or Airbnb, can hire the exact amount of employees in relation to get a profit.

Scope

This project will analyze the data for the last 20 years regarding the amount of people who visited Iceland as tourists with the purpose of predicting the visitors in the next 5 years. The data must be collected from people who entered the island by airplane or ferry. People who have visited the country with a different purpose, for instance, work, will not be considered. Due to Iceland's recent development of tourism, the data collected will be no more than 20 years ago.

Following a timeline and determined steps is crucial to reach the objective of the project in the two semesters given for it.

Problem definition

Having clear what is the problem that has to be solved is important for not doing extra work or collecting useful information.

Data collection

The data has to be collected from reliable and updated sources.

Discovery

Once we get the necessary data to reach our objective, it has to be analyzed to see any patterns it follows. EDA can be applied on this stage to make any complex data to a more understandable one.

Data Preparation

The data has to be cleaned to make an accurate prediction. Missing data must be eliminated or filled using different methods. Normalization or PCA can also be applied on this step if it is required.

Model Planning and Building

On this step, the best model for the project has to be selected. First, make clear if a supervised or unsupervised model will be needed. Once the model is trained, make sure that it is not overfitted or underfitted.

Jupyter Notebook will be used to build the model with Python language.

Results

After the model is trained, communicate the results in a presentation.

Data sources

The data that will be collected must include the number of tourists that have arrived to the country in the last 20 years, and it has to be monthly or season specified.

For this project, the data will be collected from different online sources. One of them will be Statistics Iceland. It is a website for official statistics in Iceland which has data from 1914. They collect information from different aspects which affect the country including tourism.

Another source will be [Ferðamálastofa Icelandic Tourist Board](https://www.ferdamalastofa.is/en), which is an independent authority under the Ministry of Culture and Business Affairs. One of the ways they collect the data is through surveys among domestics and foreign tourists. They are also in charge of counting the number of visitors who arrive in the international airports in Iceland and their nationalities.

Ethics

It is important to analyse any ethical consideration that can affect the result of the project.

Personal information is not required for this project, except for nationality, which may be a positive data to make sure it does not belong to just one group of people from one area.

Being aware of past events such as, economic crisis, pandemics, or natural disasters is important to understand that it can affect the results.

The government, public and private companies will be positively affected by the result of the project so that they can get a better budget spend, but it is crucial to understand that the prediction of the model will not have a 100% of accuracy.

New international treaties or visas can affect the prediction of the model.