## Abstract:

The objective of this exploratory analysis was to conduct a statistical analysis of the data of museums, aquariums and zoos. As we saw, the data set contains many null and zero values. In order to get an accurate result, we dropped the null values and got rid of the zero values from our calculations. The modeling indicates that the type of museum in each state has a positive impact on revenues.

## Design:

This project originated from T5 Data Scientists bootcamp was provided by Kaggle .This analytical report presents the most museums with higher revenues, their number in the region, and knowledge of the most famous museums for each state, as well as the most states with museums.

## Data:

The dataset is provided in .csvformat.it contains 25 columns in 33072 row all are a mix of categorical and numerical values and includes some notable columns (state , museum type , museum name , revenue ,income).

## Algorithms:

- **Data cleaning and preparation**
- checking for in missing values .
- replace missing value
- delete data that cannot be corrected /replaced
- correct any data formatting issues
- checking no duplicate

## Tools:

Python and Jupyter Notebook .
Numpy, and Pandas for Data manipulation.Matplotlib and seaborn for plotting visuializtion.

## Communication:



The most famous museums in the state

Highest state has number of museums



Counts of Muesum Type

museum type vs income or revenue