



# OPEN IIT DATA ANALYTICS

## Team : Metric Maestros

Team Leader: Zubair Bashir  
22CY10049

DATA ANALYSIS



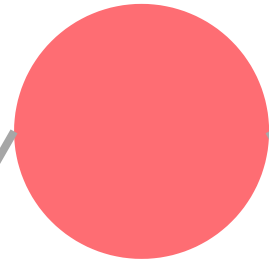
# PROBLEM STATEMENT

PREDICTING TOURIST ARRIVALS USING  
MACHINE LEARNING AND INTERNET  
SEARCH INDEX



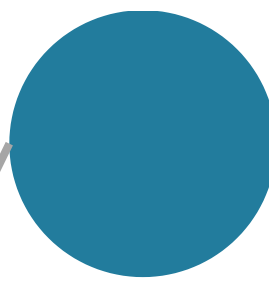
# DATA RESOURCES

**No. of holidays**



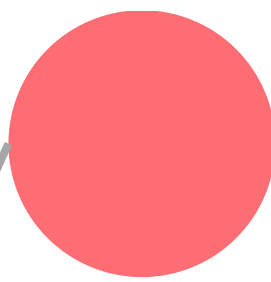
TimeAndDate.com  
GoaTourism.gov.in

**Relative humidity**



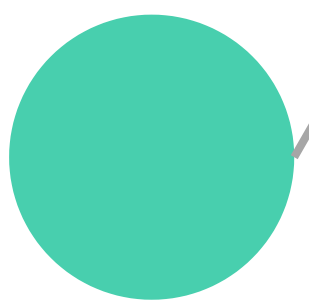
Power data  
Access Viewer

**Cultural fest**



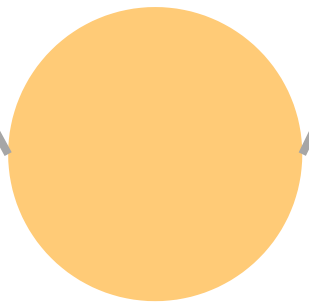
theGoaVilla

**No of Tourist**



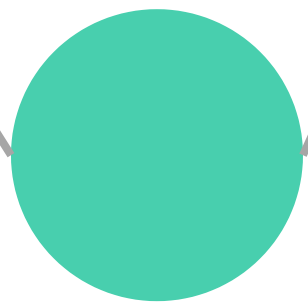
tourism.gov.in  
goaTourism.gov.in  
Google Trends

**Avg. temp.**



Google Trends  
WorldWeatherOnline

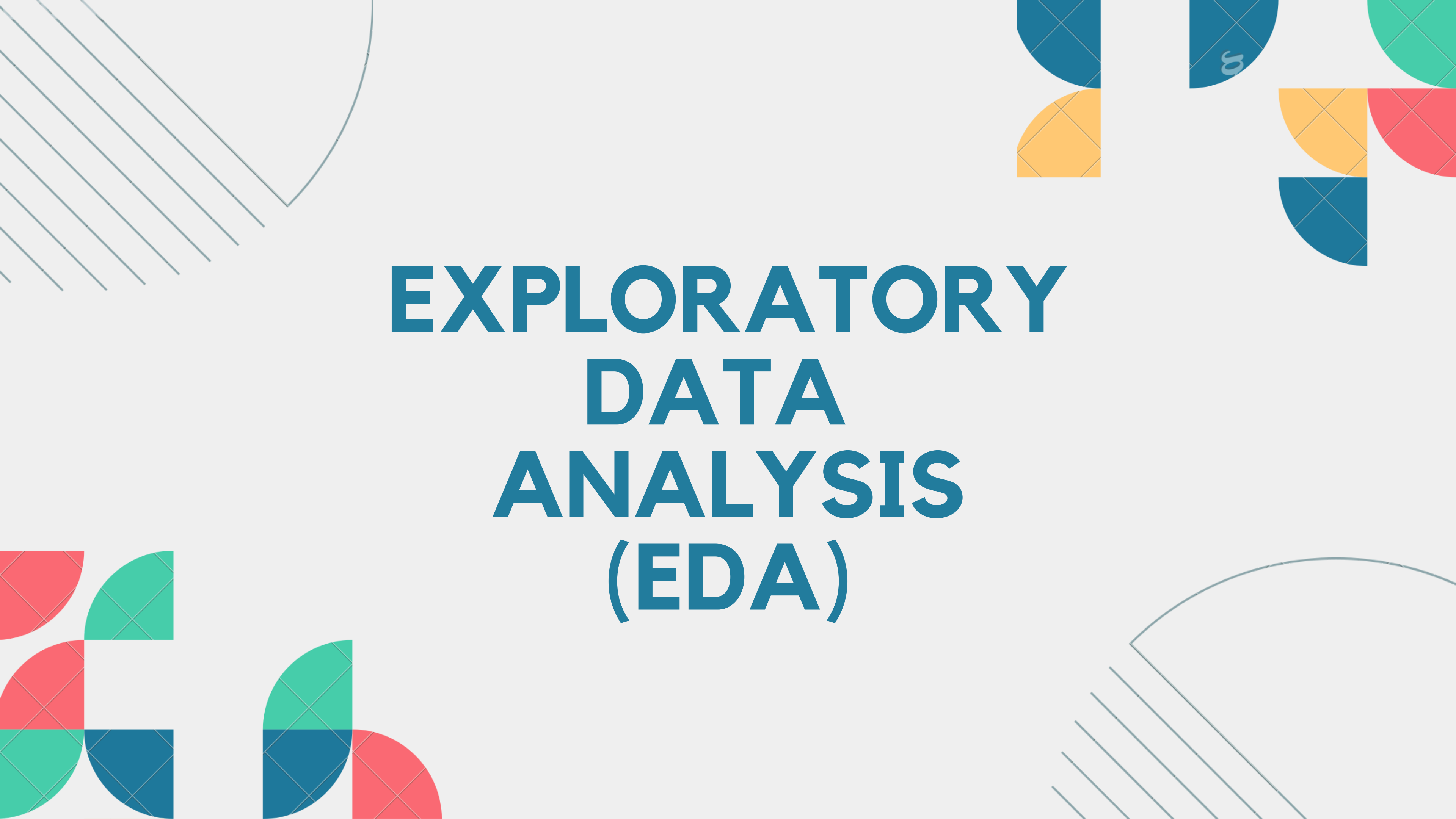
**Precipitation**



Power data  
Access Viewer

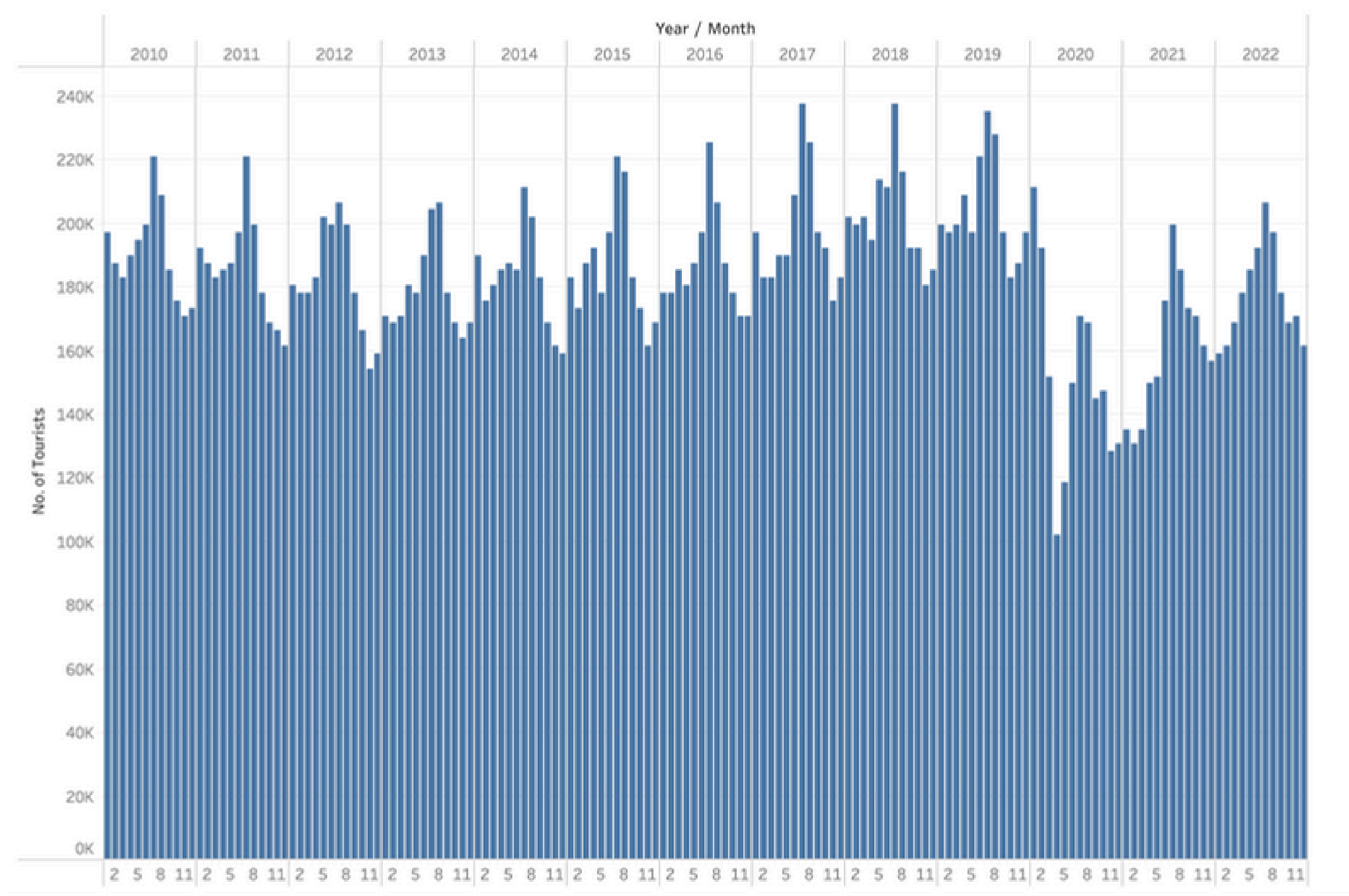
# CREATED DATA

	Year	Month	Relative	Percentage	No. of Tourists	Holidays	Avg Temp(°C)	Relative Humidity(at 2m)%	Precipitation (mm/day)	Cultural Fests	GDP of goa (k )
0	2010	1	83	0.006885	197054.6125	9	27	63.44	0.00	7	54835.000
1	2010	2	79	0.006553	187558.0046	8	27	61.88	0.00	6	54835.000
2	2010	3	77	0.006387	182809.7007	9	28	64.56	0.00	4	54835.000
3	2010	4	80	0.006636	189932.1566	7	30	70.00	0.00	5	54835.000
4	2010	5	82	0.006802	194680.4605	6	30	73.62	0.00	4	54835.000
5	2010	6	84	0.006967	199428.7644	6	27	82.56	21.09	5	54835.000
6	2010	7	93	0.007714	220796.1321	5	26	88.19	26.37	3	54835.000
7	2010	8	88	0.007299	208925.3723	7	26	89.19	15.82	5	54835.000
8	2010	9	78	0.006470	185183.8527	10	26	86.31	10.55	2	54835.000
9	2010	10	74	0.006138	175687.2449	11	26	82.44	5.27	3	54835.000
10	2010	11	72	0.005972	170938.9409	9	26	80.69	5.27	6	54835.000
11	2010	12	73	0.006055	173313.0929	11	25	69.75	0.00	7	54835.000
12	2011	1	81	0.006719	192306.3086	9	26	61.75	0.00	7	289191.045
13	2011	2	79	0.006553	187558.0046	8	26	61.75	0.00	6	289191.045
14	2011	3	77	0.006387	182809.7007	9	27	64.88	0.00	4	289191.045

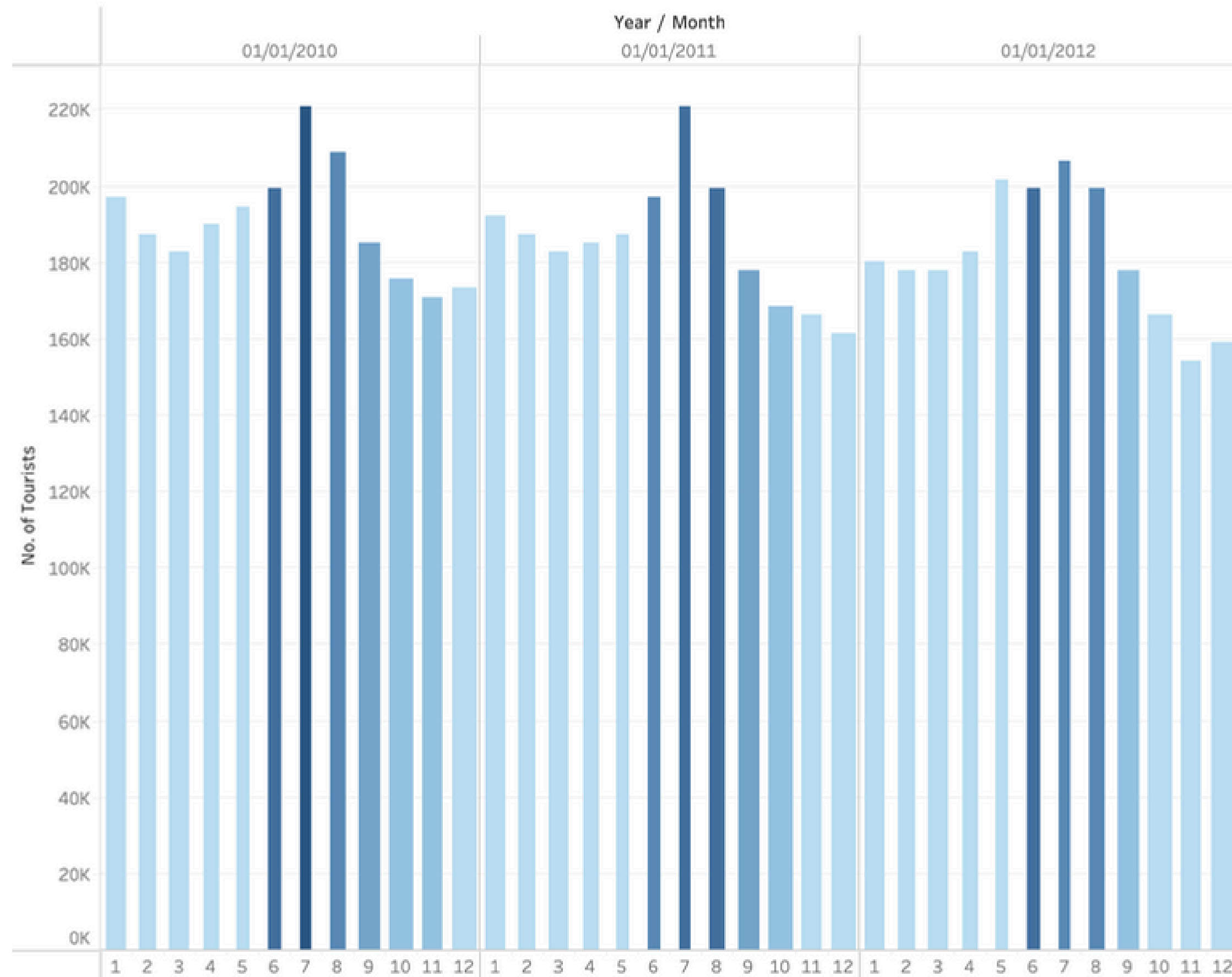
The background features decorative geometric patterns in the corners. The top-left corner has a series of parallel diagonal lines. The top-right corner contains a cluster of overlapping quarter-circles in blue, yellow, red, and green, with a small blue circle containing a white 'a' nearby. The bottom-left corner shows a 2x2 grid of quarter-circles in red, green, blue, and red. The bottom-right corner features a large quarter-circle with several parallel diagonal lines inside it.

# EXPLORATORY DATA ANALYSIS (EDA)

# No of Tourists per Month

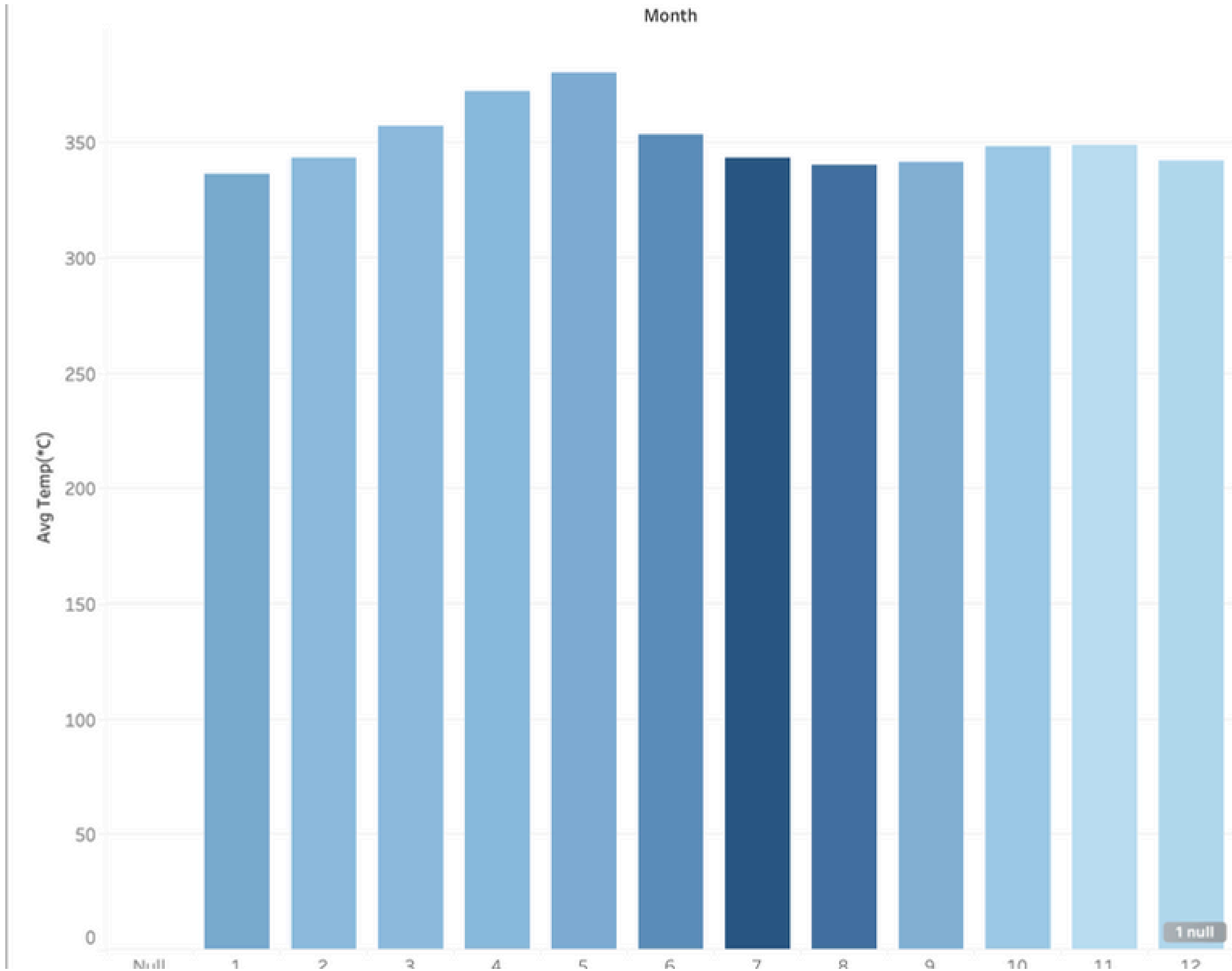


# No of Holidays and Precipitation



- Darkness: Precipitation
- Thickness: No of Holidays

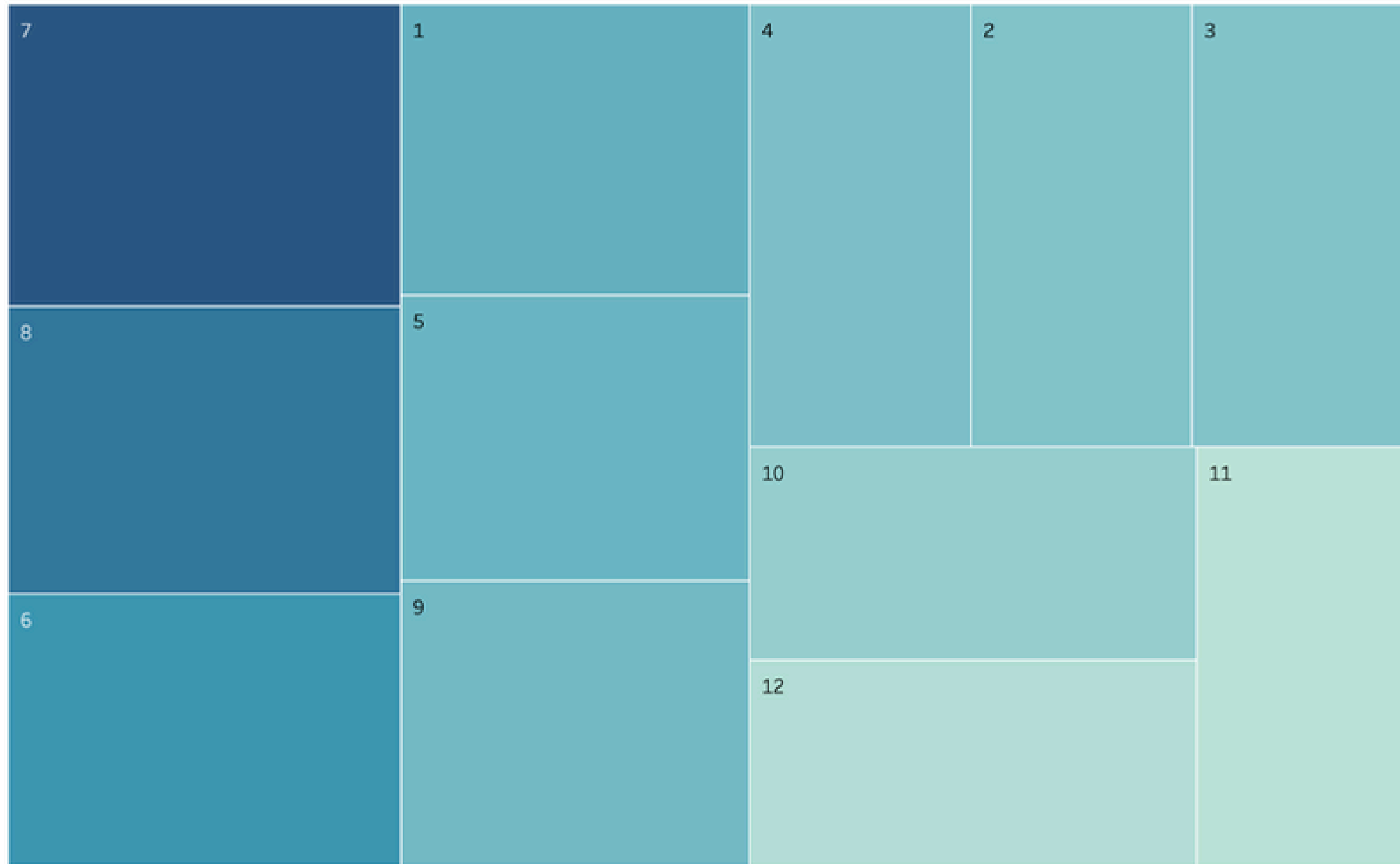
# Avg Temperature and Tourists



- Darkness: No of Tourists

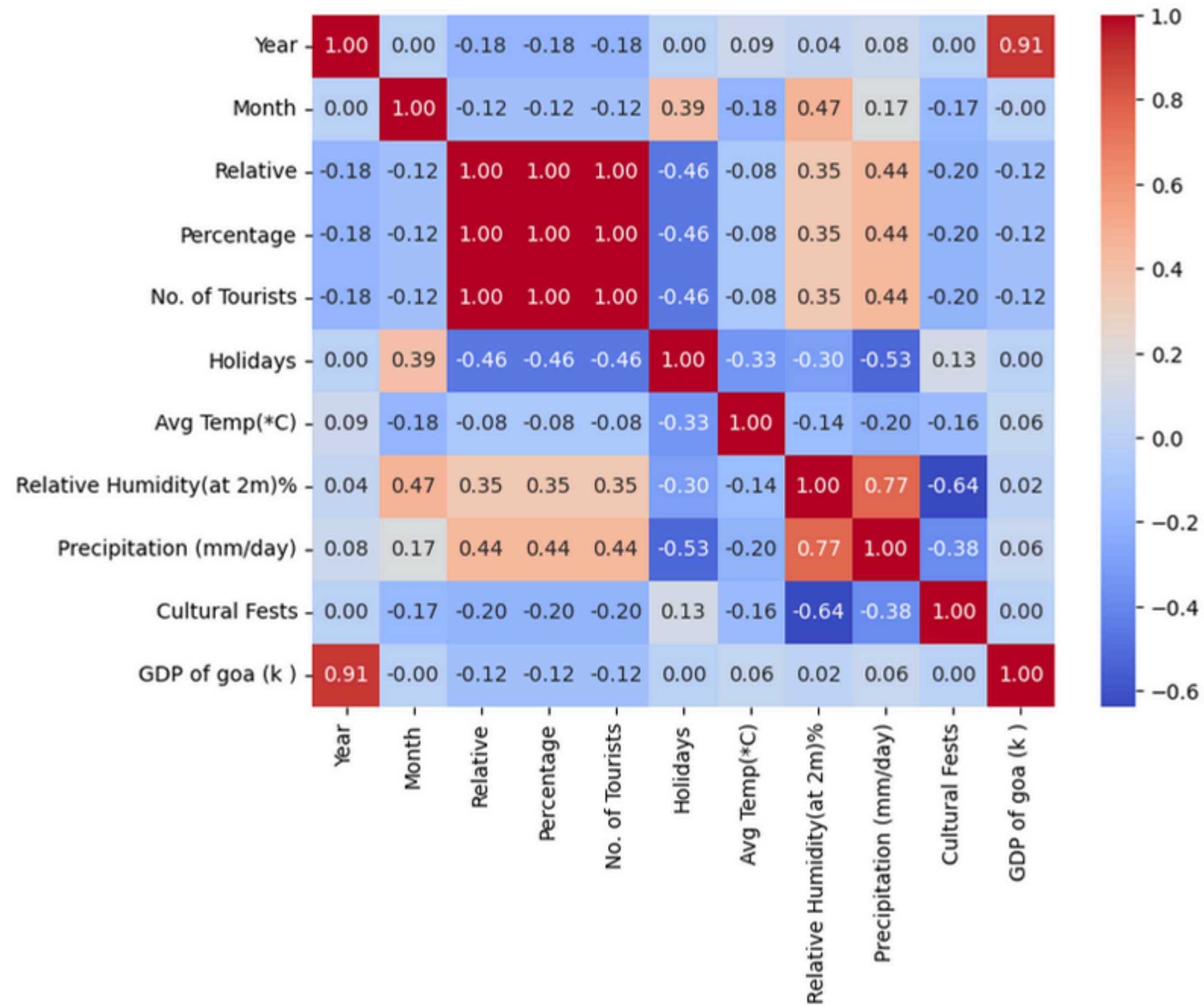


# Box Plot

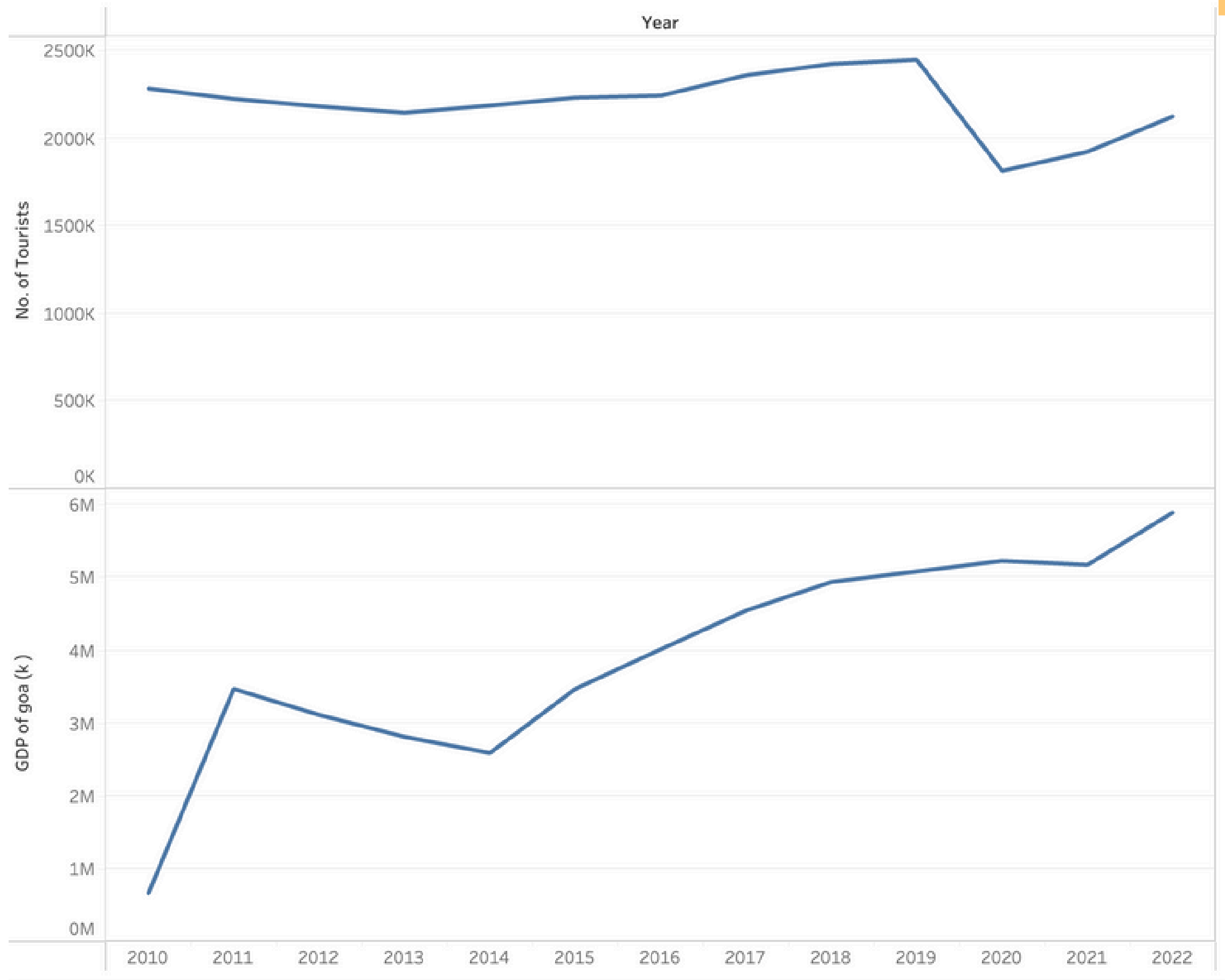


- Area: No of Tourists
- Darkness: No of Tourists

# Correlation Matrix




# Goa's Tourism and GDP





# OBSERVATION

- Tourist fluctuates monthly, with peak visitors in June-July- August.
  - People prefer visiting during more humid and low-temperature season
  - The charm of tourism rises when holidays fall.
  - No direct relation was observed between GDP and tourism
  - Even Cultural Fest doesn't seem to affect tourism much
  - During Covid, a sudden decrease in volume was observed but still, the trend remains the same.
- 



ML MODELS


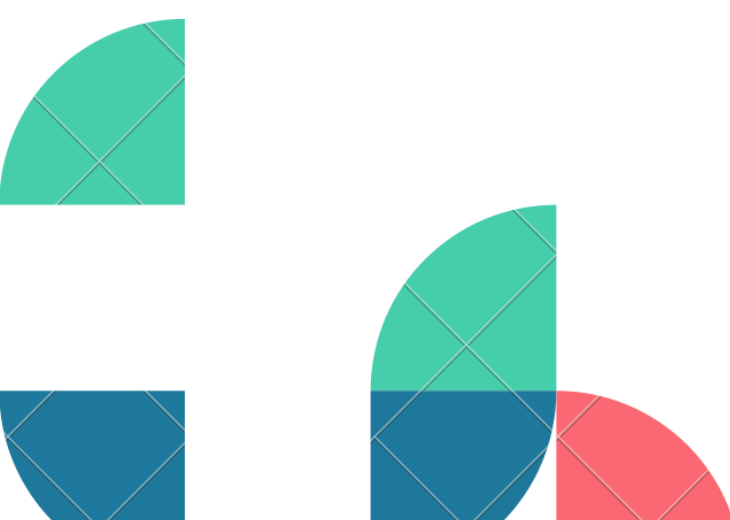
VECTOR AUTOREGRESSIVE  
AND  
PROPHET MODEL





# Vector Autoregressive (VAR)

Vector Autoregressive is a statistical and time-series analysis model used for analyzing and forecasting multivariate time series data. It's a type of autoregressive model that deals with multiple time series variables simultaneously.





# Prophet Model

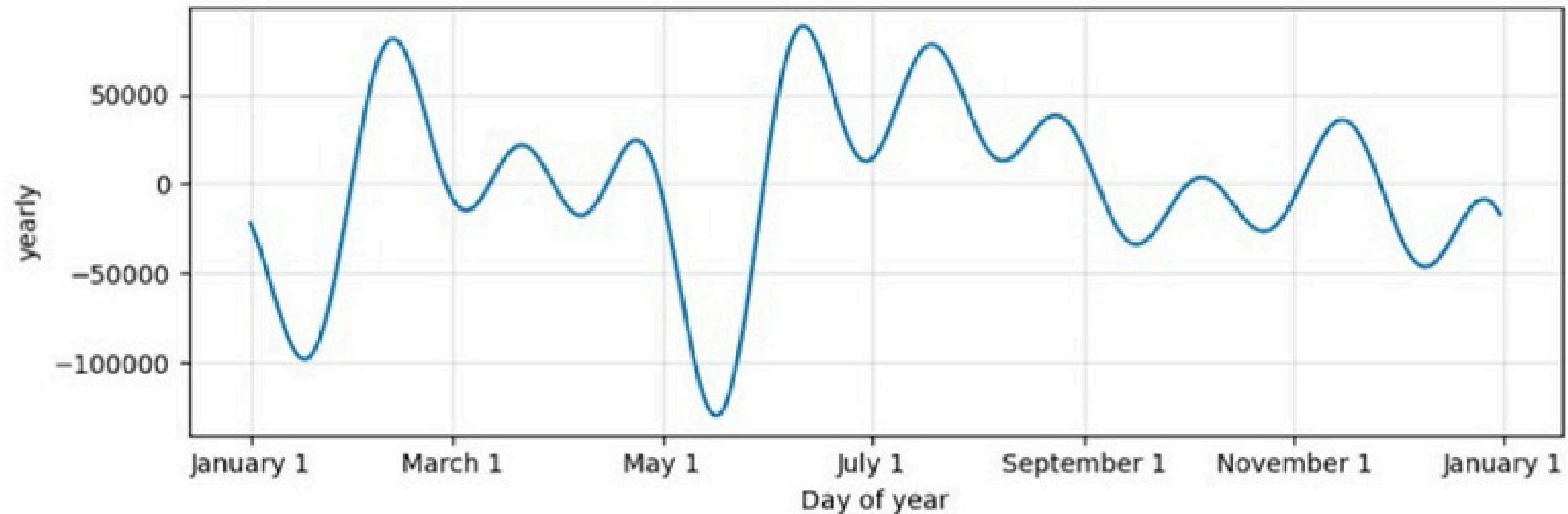
The Prophet model is a forecasting tool developed by Facebook's Core Data Science team, designed to make time series forecasting more accessible and straightforward for non-experts. It is particularly well-suited for applications in business and economics, where historical data is available and seasonality and holidays play a significant role in forecasting.





# Prophet Model

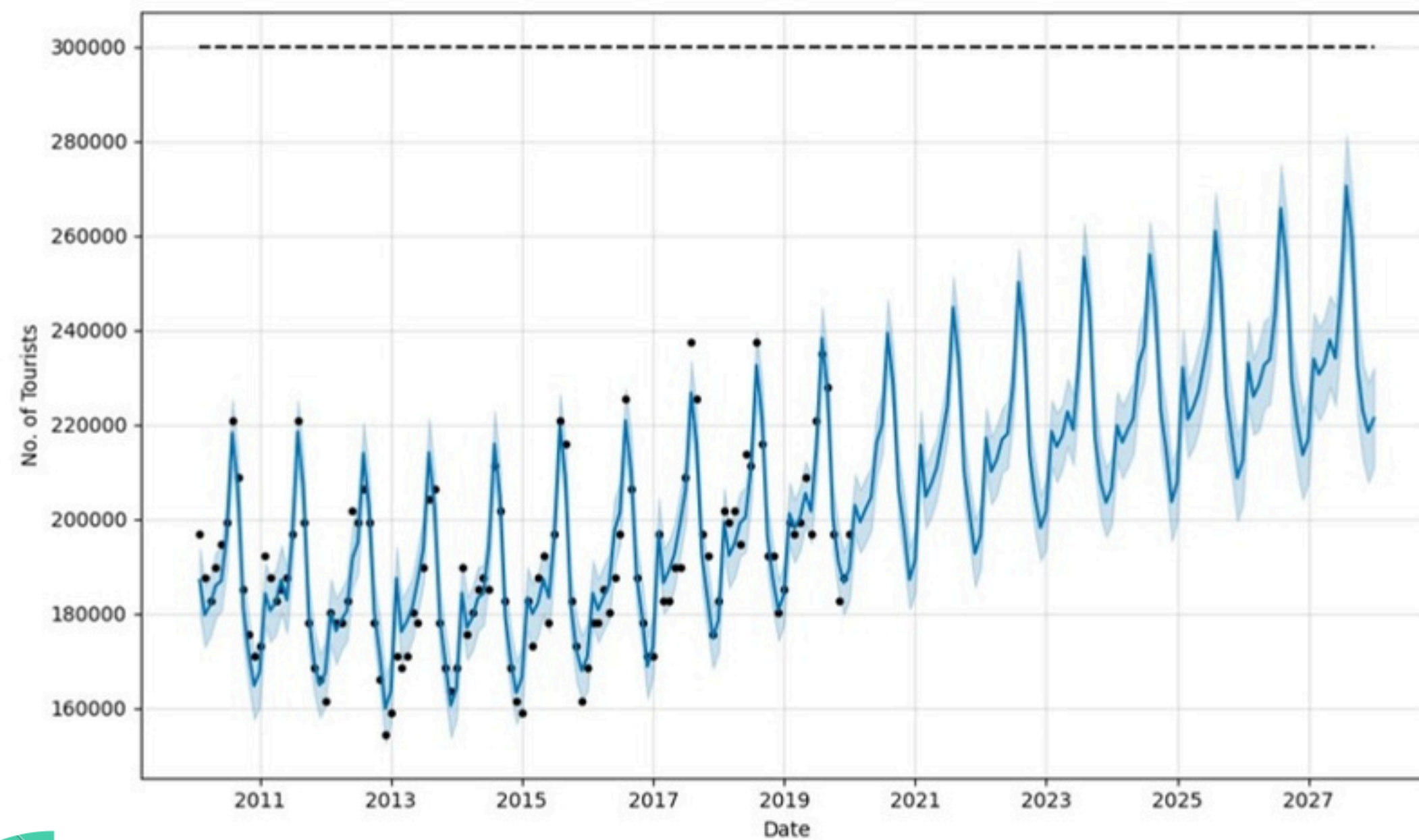
## Monthly Trend Analysis





# Prophet Model

## Yearly Trend Analysis



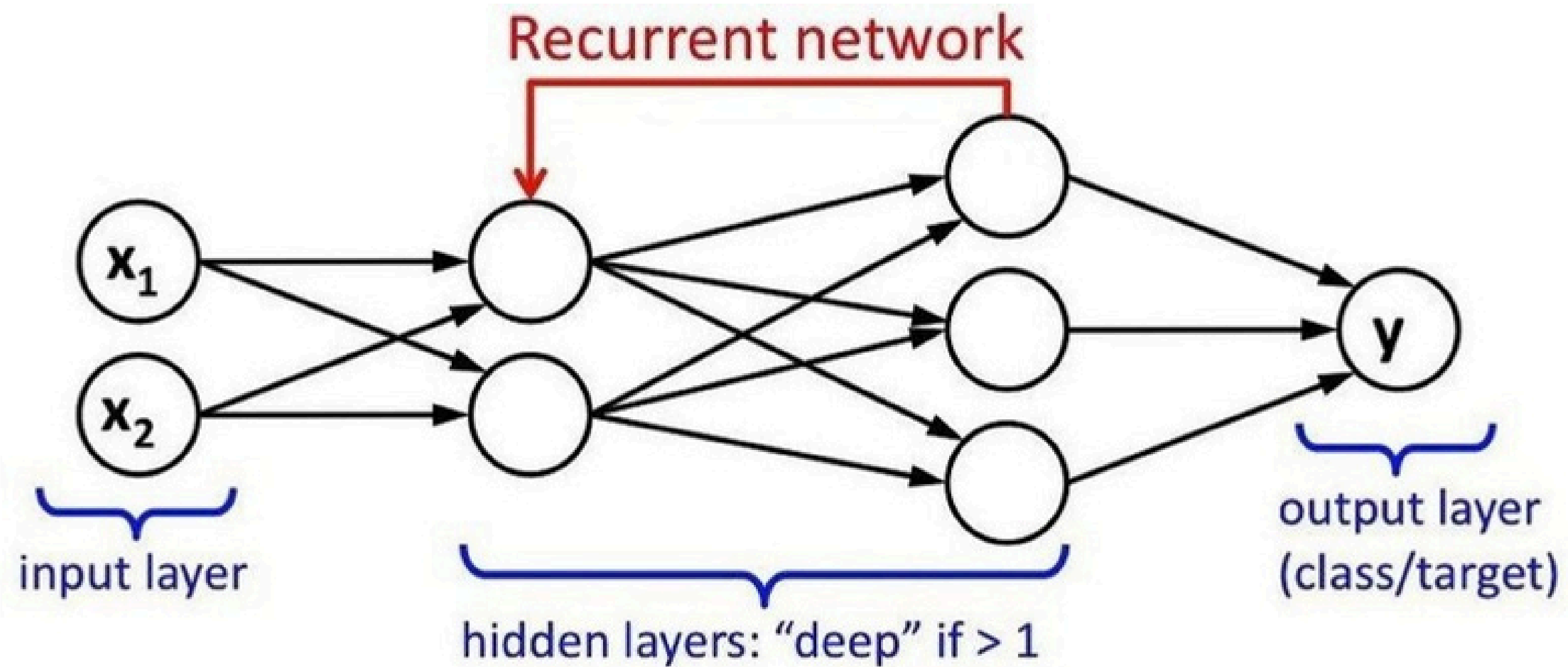


**DEEP LEARNING MODEL**

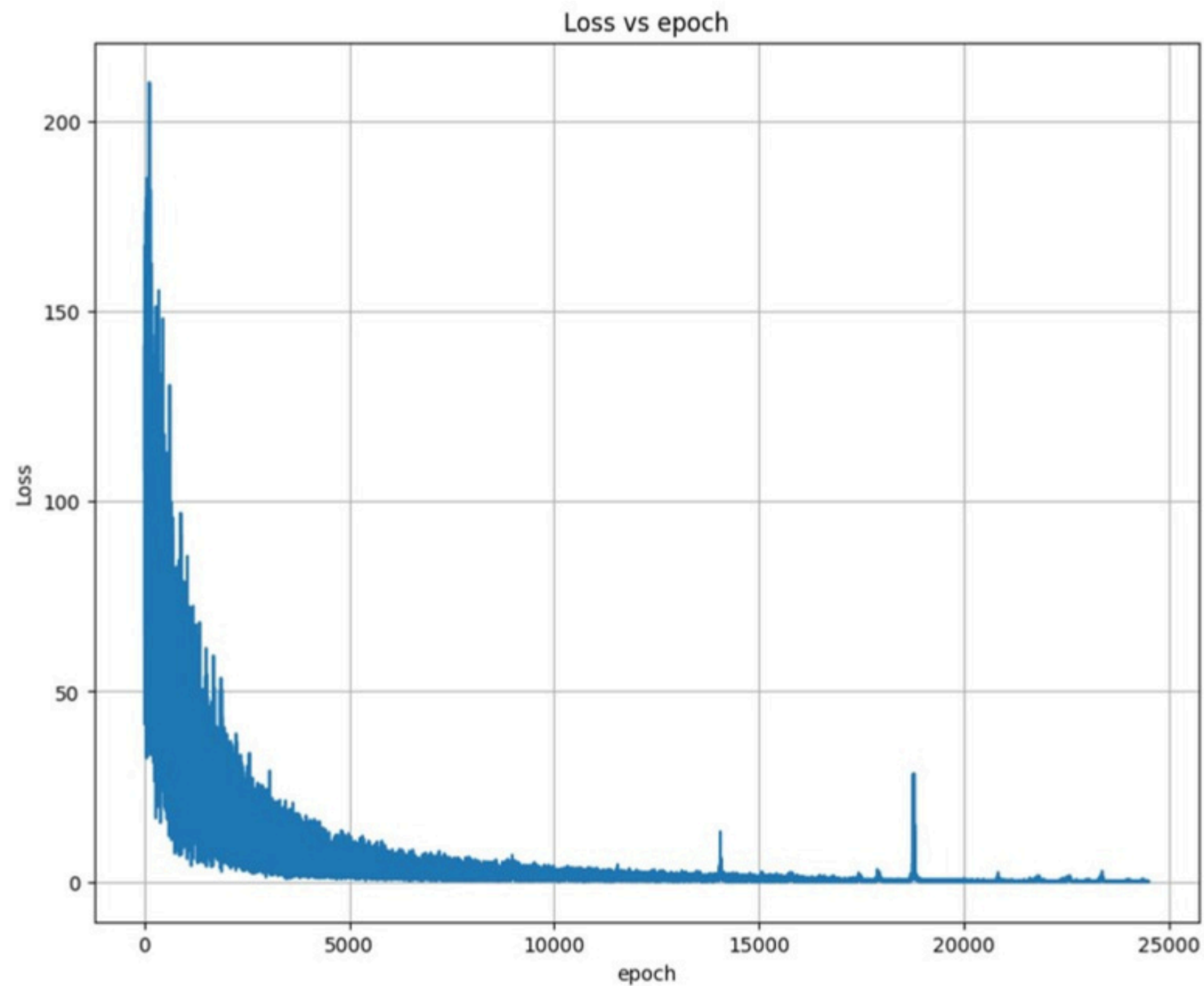
**RECURRENT NEURAL  
NETWORK**



# Our Model



# Training Loss vs Epoch



# RESULTS

Model	MSE	Accuracy
VAR (monthly)	2167.44	97.83
VAR (yearly)	16556.51	83.44
Prophet (monthly)	9528.22	94.45
Prophet (yearly)	10045.56	90.54
RNN model	0.34	95.67

MSE -> Mean Squared Error  
MAPE -> Mean Absolute Percentage Error

The background features several decorative geometric patterns. In the top-left corner, there are thin, parallel diagonal lines. In the top-right corner, there is a cluster of overlapping semi-circles in blue, green, and red. In the bottom-left corner, there is another cluster of overlapping semi-circles in blue, green, and red, with a small blue semi-circle containing a white lowercase 'v'. In the bottom-right corner, there is a large, faint, light blue arc and some thin diagonal lines.

**THANK YOU**