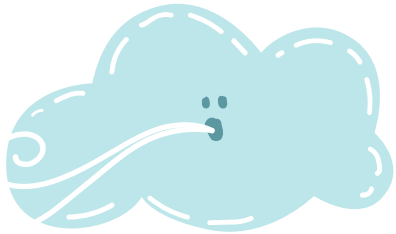
Project 1 :

Data Analysis of Singapore Rainfall

Oct 17, 2023 : DSI 03

Agenda

- Problem statement
- Data Process
- Data Dictionary
- Data Overview
- Actionable Insights
- Summary and Recommendation



Problem Statement

“ CEO of Otteri Wash & Dry, a self-service laundry, in Singapore branch wanted to optimize company's business in term of revenue gathering and cost-saving to rely on local climate conditions. So he started to explore monthly climate statistics from Jan 1984 to Aug 2022 in order to initiate annual strategic planning. ”

Data Process

1. Merge DataFrames by using merge (left) with `'month'` as primary key between 1982-Jan and 2022-Aug.
2. Extract `'month'` into 2 new columns : `'year'` and `'month_name'`
3. Add a new column to DataFrame - `'monsoon_type'`
4. Drop data between 1982-Jan and 1983-Dec as the climate station was relocated. (24 observations were eliminated, 464 observations remained)

Data Dictionary

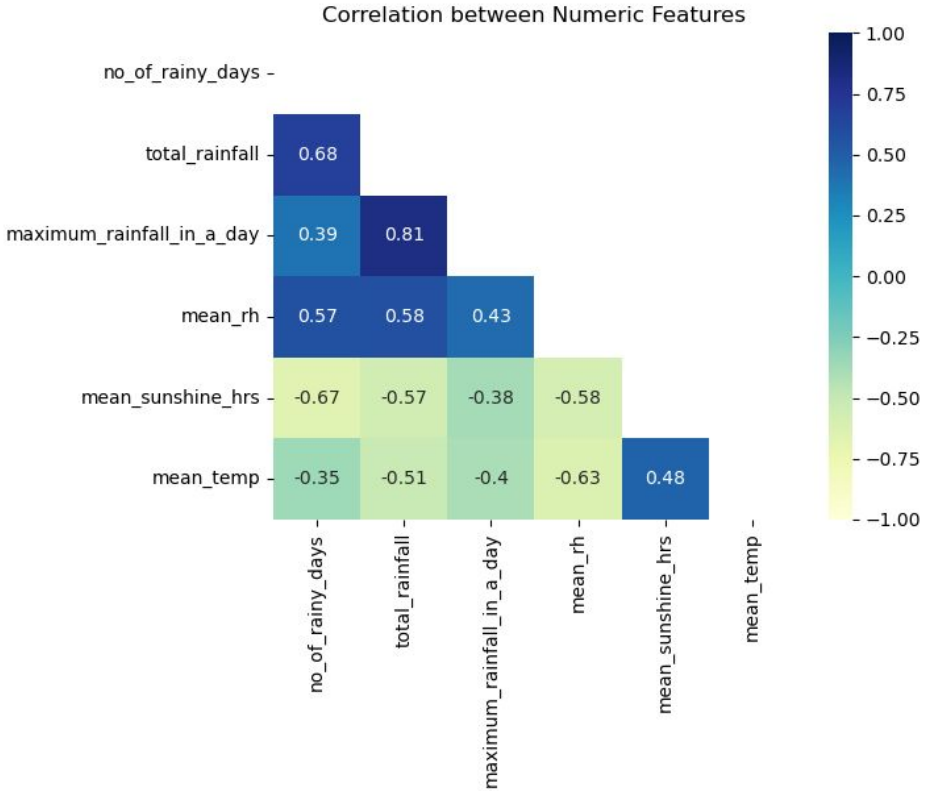
Feature	Type	Dataset	Description
month	string	rainfall-monthly-number-of-rain-days	Datetime (Month) "YYYY-MM"
no_of_rainy_days	int	rainfall-monthly-number-of-rain-days	Number of Rain Days in the Month
total_rainfall	float	rainfall-monthly-total	Monthly Total Rainfall(Millimetre)
maximum_rainfall_in_a_day	float	RainfallMonthlyHighestDailyTotal	Highest Daily Rainfall in the Month(Millimetre)
mean_rh	float	RelativeHumidityMonthlyMean	Monthly mean relative humidity(%)
mean_sunshine_hrs	float	SunshineDurationMonthlyMeanDailyDuration	Monthly Mean Daily Sunshine Duration(Hours)
mean_temp	float	SurfaceAirTemperatureMonthlyMean	Surface Air Temperature - Monthly Mean(Degree Celsius)
temp_mean_daily_min	float	SurfaceAirTemperatureMonthlyMeanDailyMinimum	Monthly Mean Daily Minimum Temperature(Degree Celsius)
temp_mean_daily_max	float	SurfaceAirTemperatureMonthlyMeanDailyMaximum	Monthly Mean Daily Maxumum Temperature(Degree Celsius)
month_name	string	Added column	Name of the month e.g. Jan, Feb
year	int	Added column	Year (B.C.)
moonsoon_type	string	Added column	To distinguish the type of moonsoon season by month (NE,SW,None)

Data Overview

```
RangeIndex: 464 entries, 0 to 463
Data columns (total 12 columns):
 #   Column                      Non-Null Count  Dtype
---  ------                      --------------  -----
 0   month                       464 non-null    datetime64[ns]
 1   no_of_rainy_days            464 non-null    int64
 2   total_rainfall              464 non-null    float64
 3   maximum_rainfall_in_a_day   464 non-null    float64
 4   mean_rh                     464 non-null    float64
 5   mean_sunshine_hrs           464 non-null    float64
 6   mean_temp                   464 non-null    float64
 7   temp_mean_daily_min         464 non-null    float64
 8   temp_mean_daily_max         464 non-null    float64
 9   year                        464 non-null    int64
 10  month_name                  464 non-null    object
 11  moonsoon_type               464 non-null    object
dtypes: datetime64[ns](1), float64(7), int64(2), object(2)
```

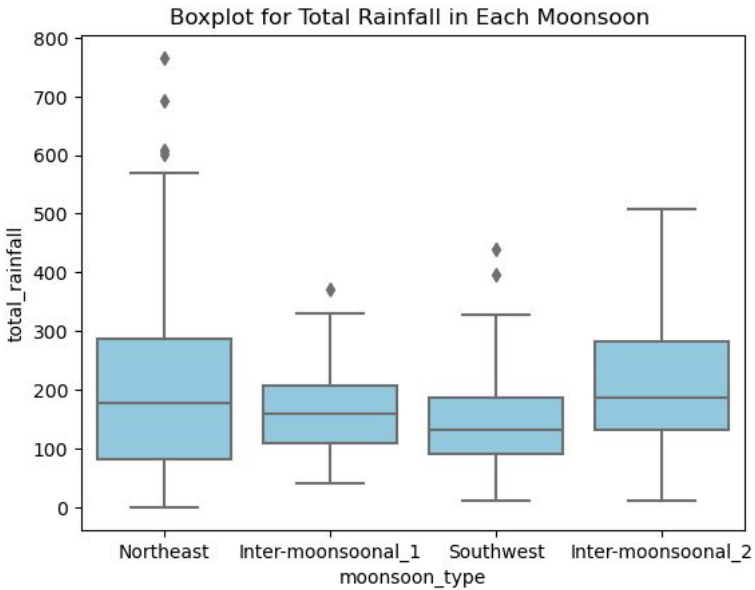
	no_of_rainy_days	total_rainfall	maximum_rainfall_in_a_day	mean_rh	mean_sunshine_hrs	mean_temp	temp_mean_daily_min	temp_mean_daily_max
count	464.000000	464.000000	464.000000	464.000000	464.000000	464.000000	464.000000	464.000000
mean	14.196121	178.621767	52.555172	82.253879	5.688362	27.687500	24.929095	31.525431
std	4.898793	113.731148	35.274615	3.429479	1.180362	0.794484	0.731839	0.844164
min	1.000000	0.200000	0.200000	72.000000	3.000000	25.500000	22.900000	28.800000
25%	11.000000	96.550000	31.000000	79.975000	4.800000	27.100000	24.400000	31.000000
50%	14.000000	159.700000	44.200000	82.700000	5.700000	27.700000	24.900000	31.500000
75%	18.000000	239.650000	63.425000	84.725000	6.500000	28.300000	25.400000	32.100000
max	27.000000	765.900000	216.200000	90.700000	9.200000	29.500000	27.100000	34.100000

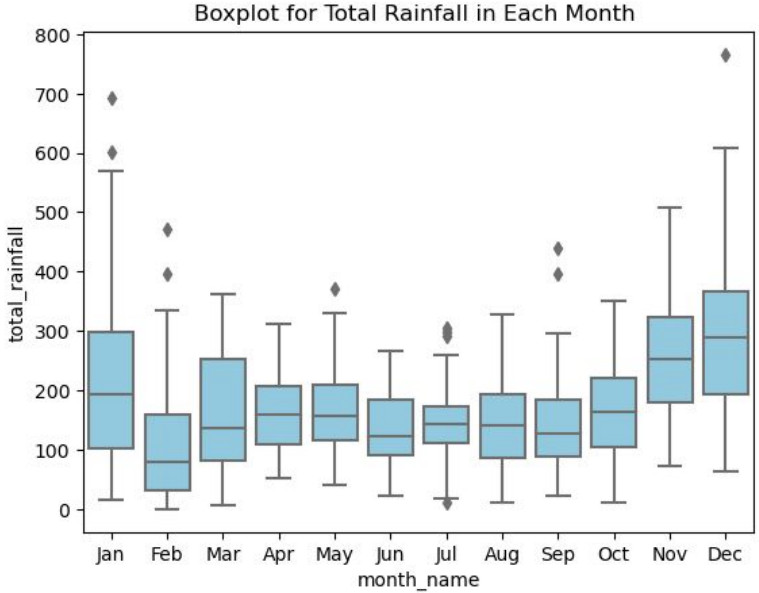
‘maximum_rainfall_in_a_day’ and ‘total_rainfall’ has the strongest positive correlation while ‘mean_sunshine_hrs’ and ‘no_of_rainy_days’ has the strongest negative correlation.



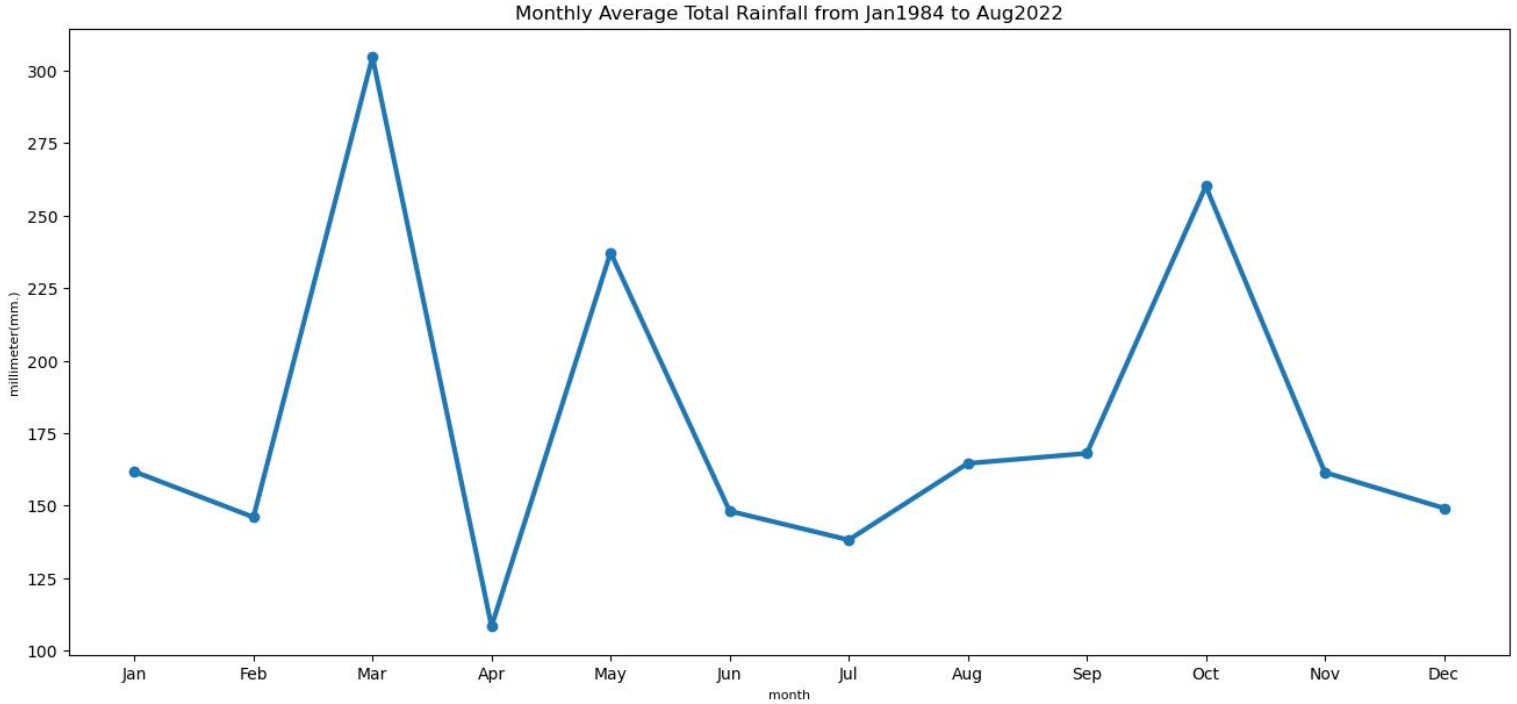
Remark :
Northeast Monsoon Monsoon : December - March
Southwest Monsoon monsoon : June - September
Inter monsoonal 1 : April - May
Inter monsoonal 2 : October - November

Northeast Monsoon and Inter-monsoonal 2 seems to have higher rainfall when comparing with the other two.



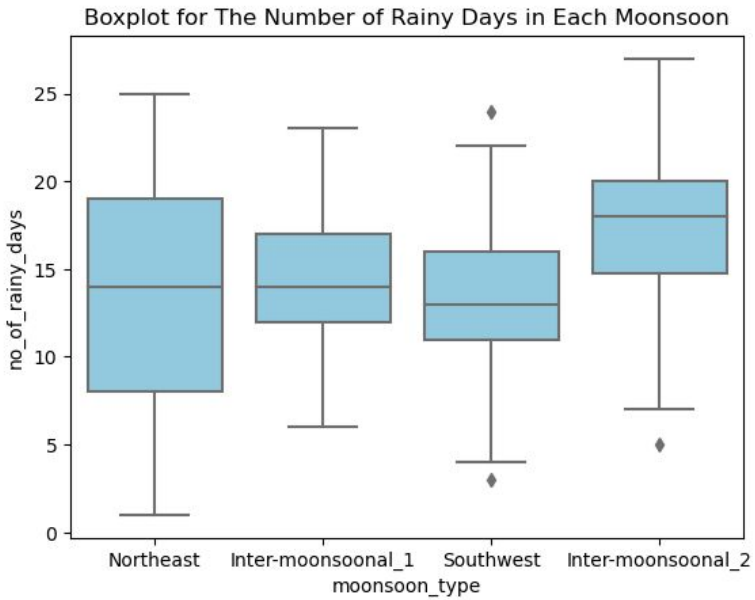


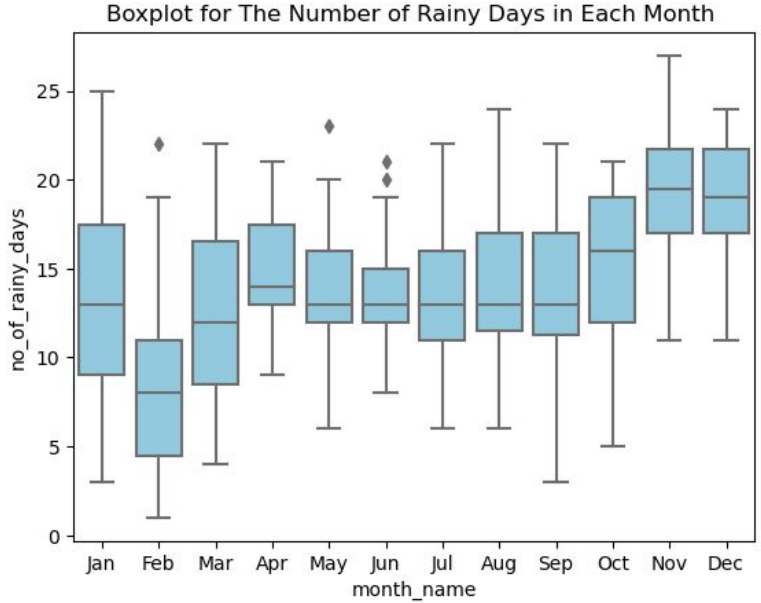
Top 2 months having highest rainfall are in Northeast Monsoon and Inter-monsoonal 2 respectively.



Remark :
Northeast Monsoon Monsoon : December - March
Southwest Monsoon monsoon : June - September
Inter monsoonal 1 : April - May
Inter monsoonal 2 : October - November

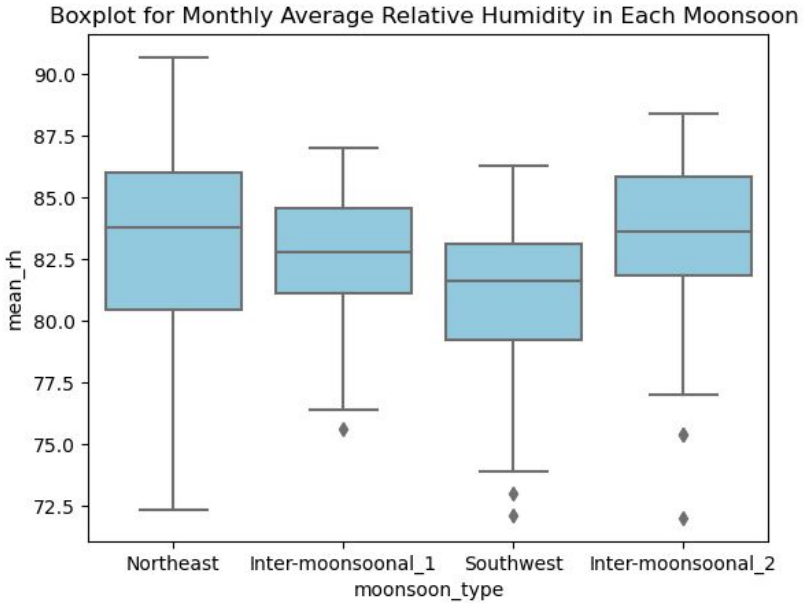
Northeast Monsoon and Inter-monsoonal 2 seems to have more number of rainy days when comparing with the other two.

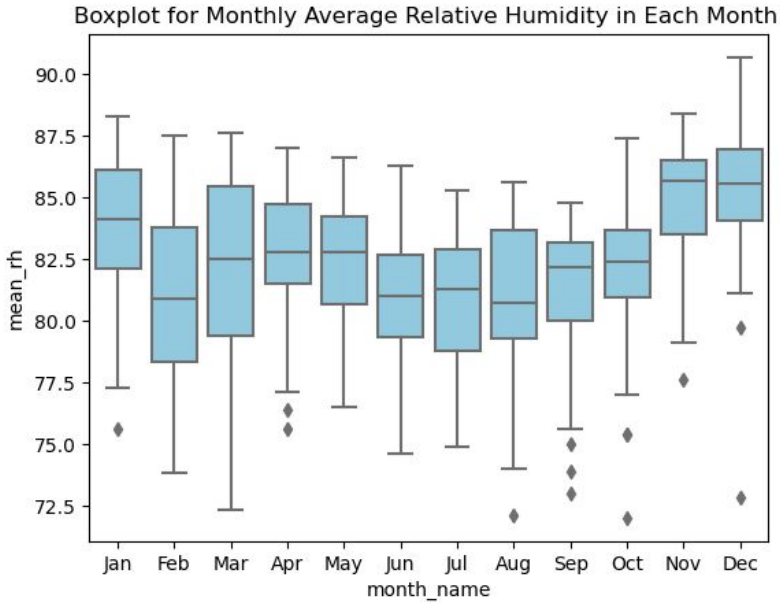




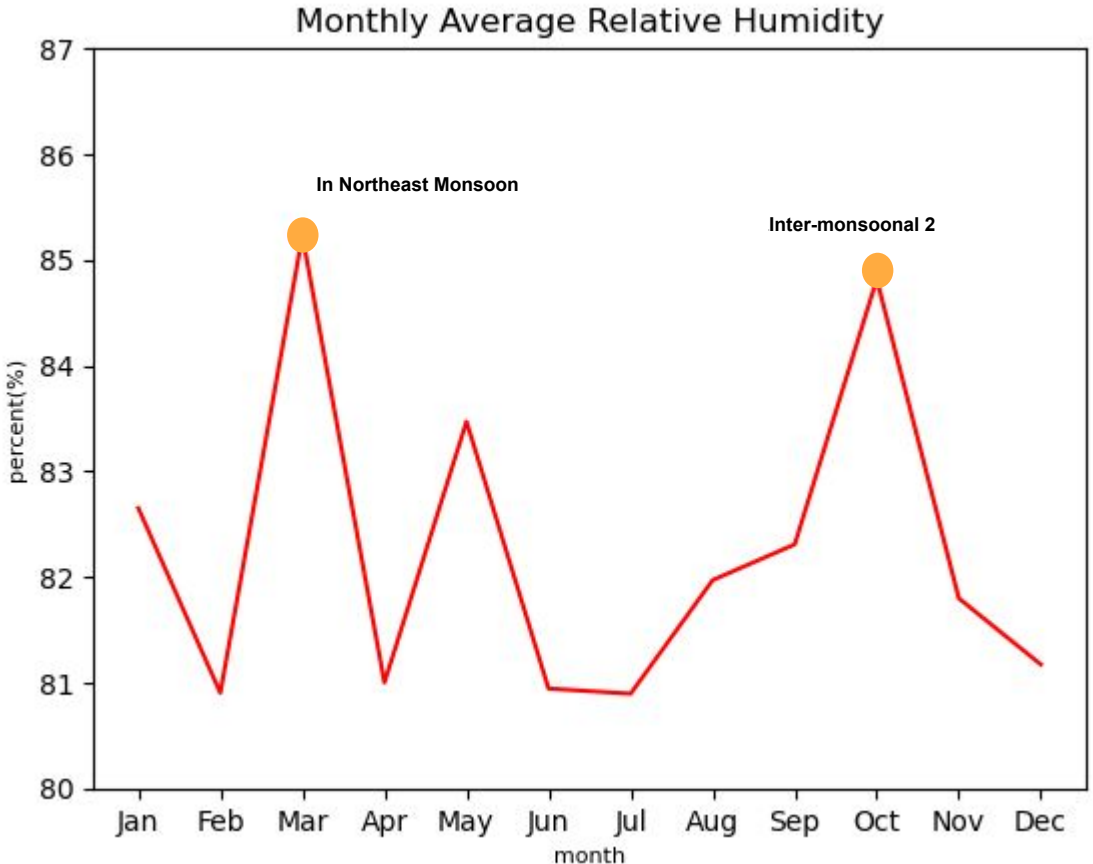
Remark :
Northeast Monsoon Monsoon : December - March
Southwest Monsoon monsoon : June - September
Inter monsoonal 1 : April - May
Inter monsoonal 2 : October - November

Northeast Monsoon and Inter-monsoonal 2 seems to have higher relative humidity when comparing with the other two.



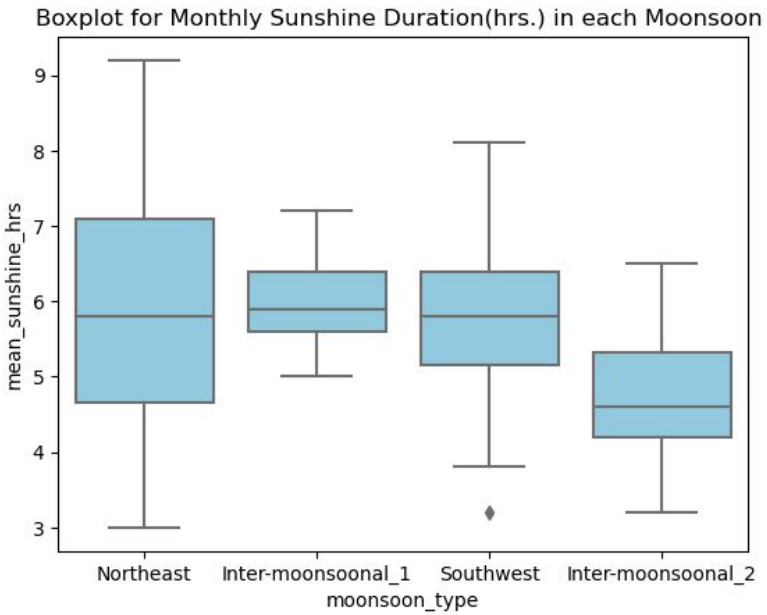


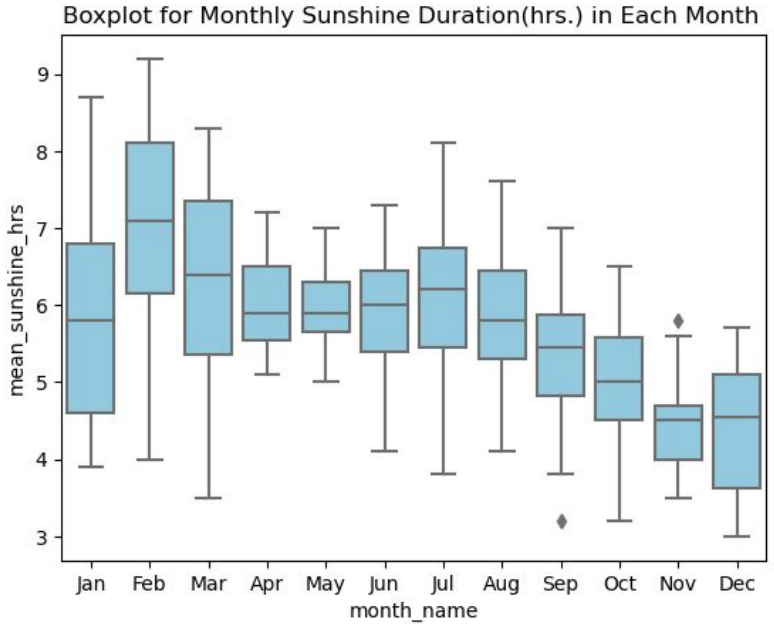
Top 2 wettest months are in Northeast Monsoon and Inter-monsoonal 2 respectively.



Remark :
Northeast Monsoon Monsoon : December - March
Southwest Monsoon monsoon : June - September
Inter monsoonal 1 : April - May
Inter monsoonal 2 : October - November

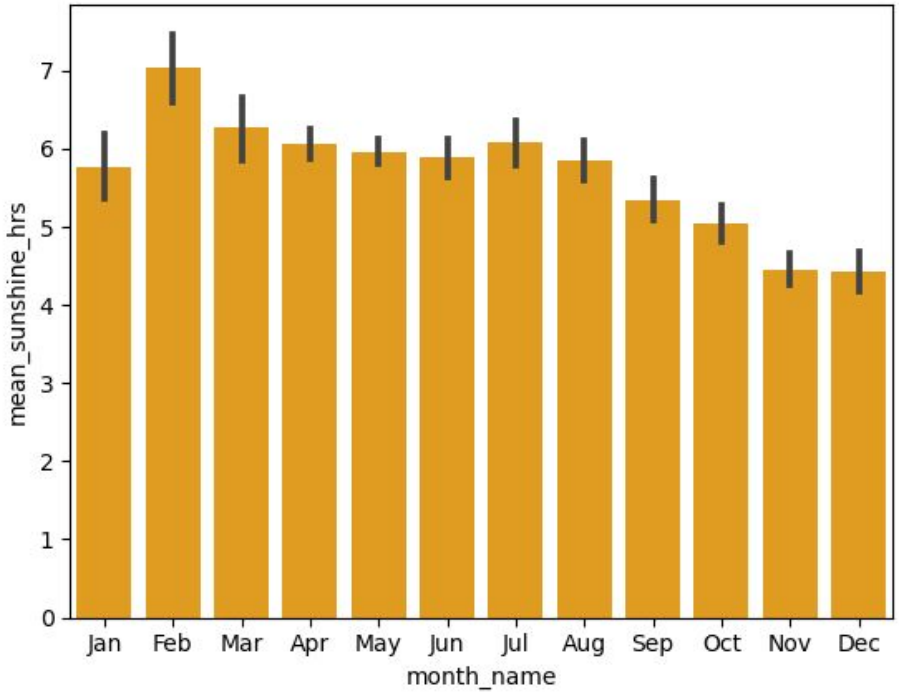
Northeast Monsoon and Southwest Monsoon turn to have higher hours of sunshine duration when comparing with the other two





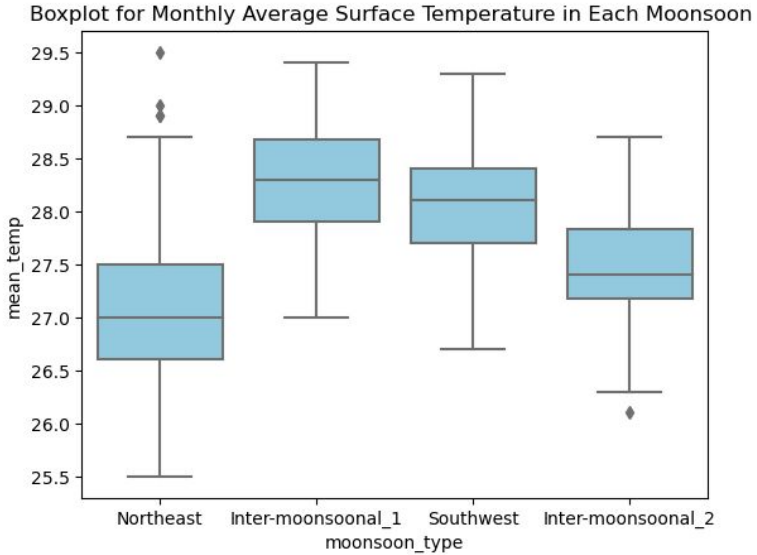
Remark :
Northeast Monsoon Monsoon : December - March
Southwest Monsoon monsoon : June - September
Inter monsoonal 1 : April - May
Inter monsoonal 2 : October - November

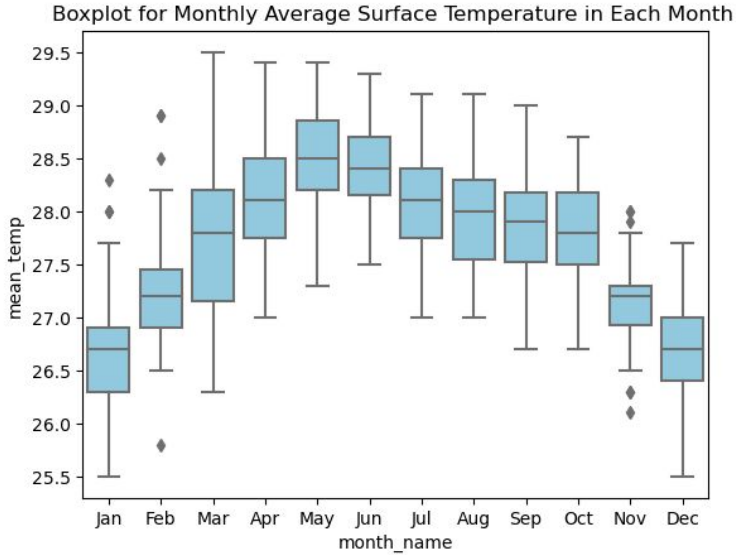
High sunshine duration is between February and August



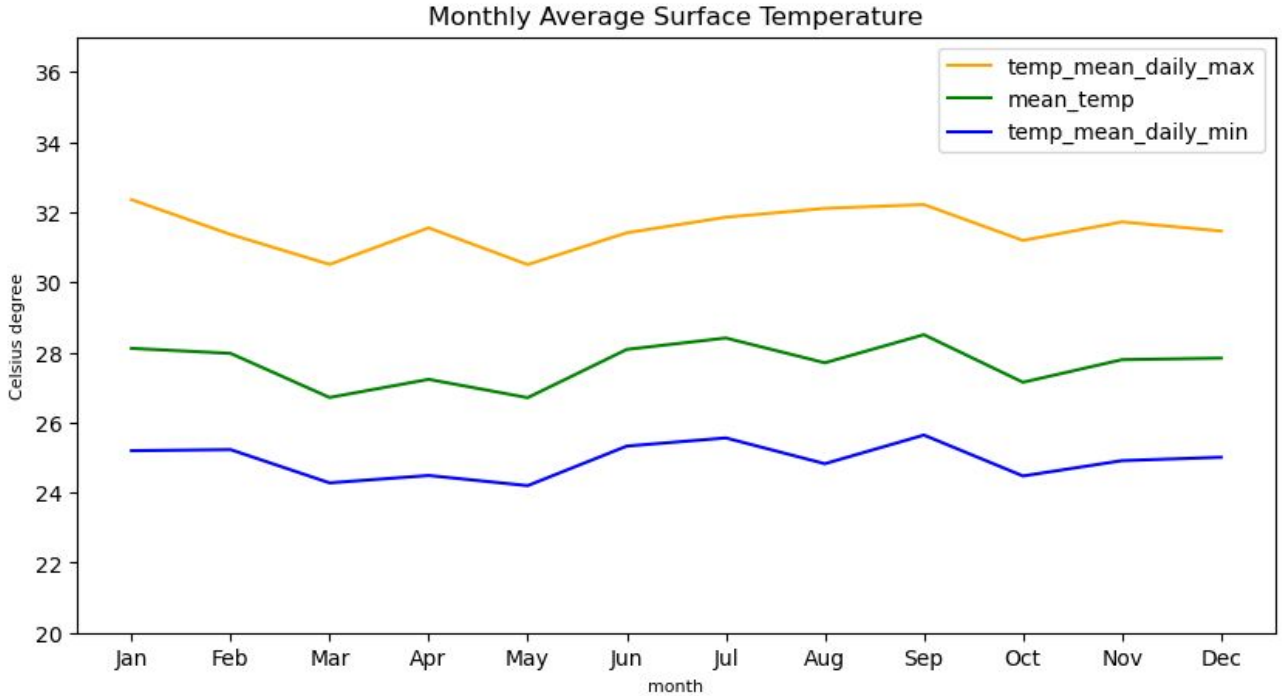
Remark :
Northeast Monsoon Monsoon : December - March
Southwest Monsoon monsoon : June - September
Inter monsoonal 1 : April - May
Inter monsoonal 2 : October - November

Northeast Monsoon and Southwest Monsoon turn to be hotter when comparing with the other two.

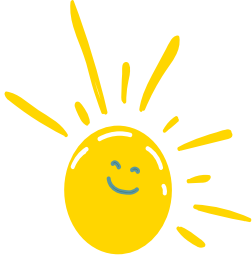




Between Inter monsoon 1 and Southeast monsoon, the surface temperature slightly increased and dropped in October. (End of Summer Time)



Summary & Recommendation



By the analysis, for Northeast Monsoon and inter_monsoonal 2 are the wettest period ,The CEO should go for ‘forward-strategy’ in order to gather revenues and new customers as the high rainfall and humidity, people might need an instant solution when drying clothes themselves is difficulty,

while for Southwest Monsoon and inter_monsoonal 1, The CEO should aim to ‘backward-strategy’ in order to maintain or decrease the cost and retain customers with CRM strategies as lower rainfall and high temperature with longer sunshine duration which can give convenience when drying clothes outdoor.