Climate Factors Forecast

Group Members: Goel Ishita,Khoo Chun Yun,ShuWei Ang,Yu Qing,Zhao Jiayi

Overview: Climate Change in the World and in Singapore

In the world:

- Long-lasting change of the temperature and weather patterns.
- Many human activities contribute to climate change.
- Has become a very challenging global issue.

In Singapore:

- Experiencing more frequent extreme weather.
- Impacted Singapore's temperature and precipitation patterns.

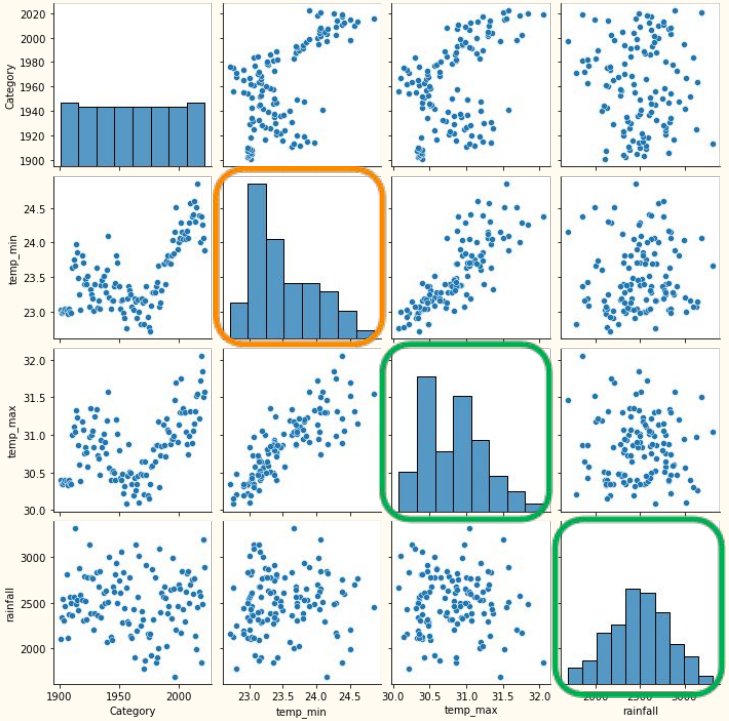
Introduction: Why temperature and rainfall in Singapore?

- Singapore is located near the equator, very warm all year round.
- Singapore is also close to the ocean, with frequent precipitation.
- Due to these features, we wish to study the changes in precipitation, maximum temperature and minimum temperature in Singapore from 1901 to 2022.
- Temperature changes are the main manifestation of climate change
- Climate change can also lead to changes in precipitation patterns.

Data Description

Variables	description	type	e.g.
Category	Year	integer	1901, 1902,..., 2022
temp_min	Minimum temperature for the year	numeric (float)	23.03, 22.98,..., 23.89
temp_max	Maximum temperature for the year	numeric (float)	30.4, 30.34,…,31.58
rainfall	Total amount of rain collected for the year (mm)	numeric (float)	2106.77, 2334.47,..., 2894.53

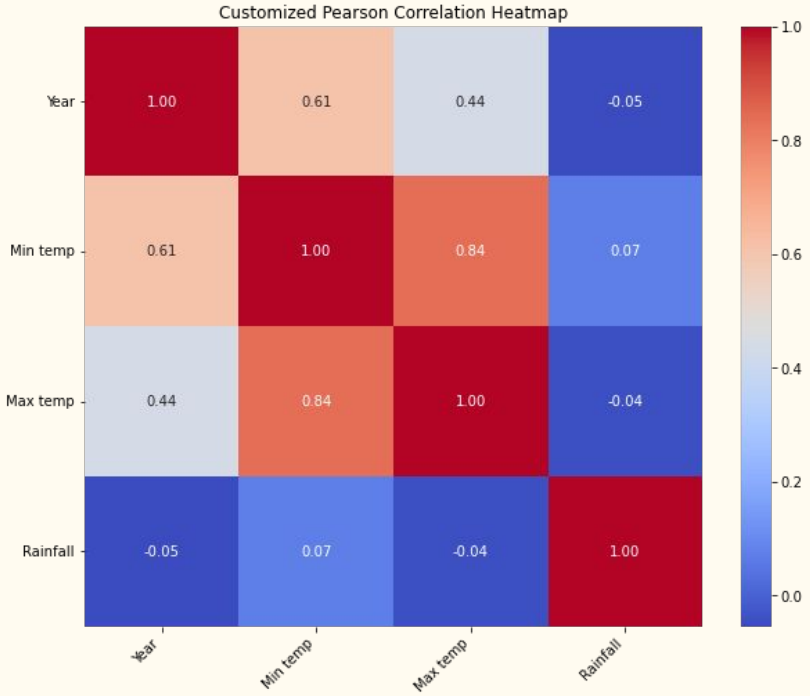
Data Exploration



Distribution:

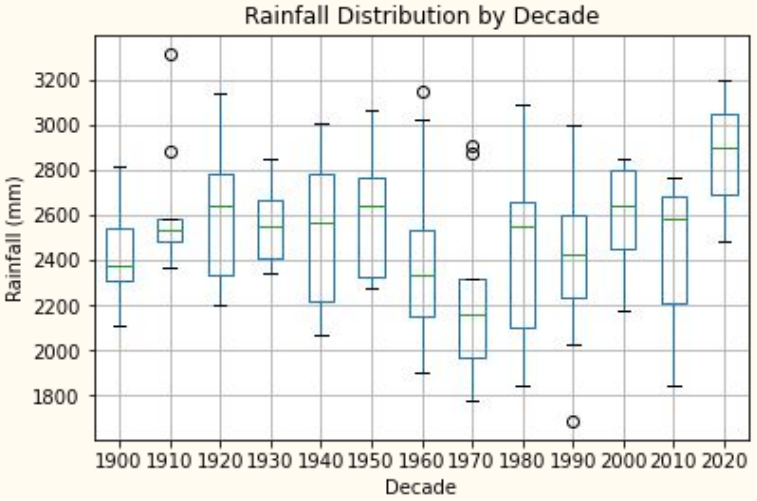
- Minimum temperature is more right-skewed.
- Maximum temperature and rainfall resemble more of a normal distribution.

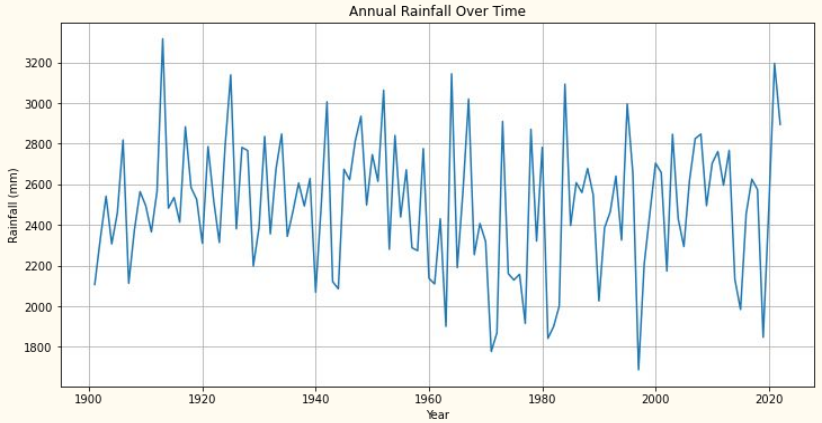
Data Exploration



- Strong correlation between minimum temperature and maximum temperature
- Relatively strong correlation can also be observed between rainfall and minimum temperature
- Weak/no correlation between rainfall and the other variables

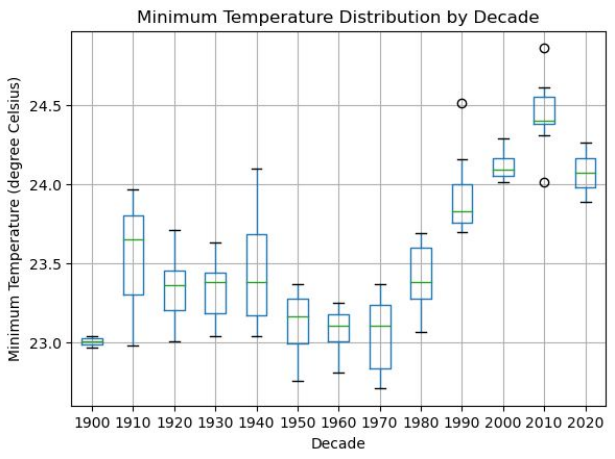
Data Exploration

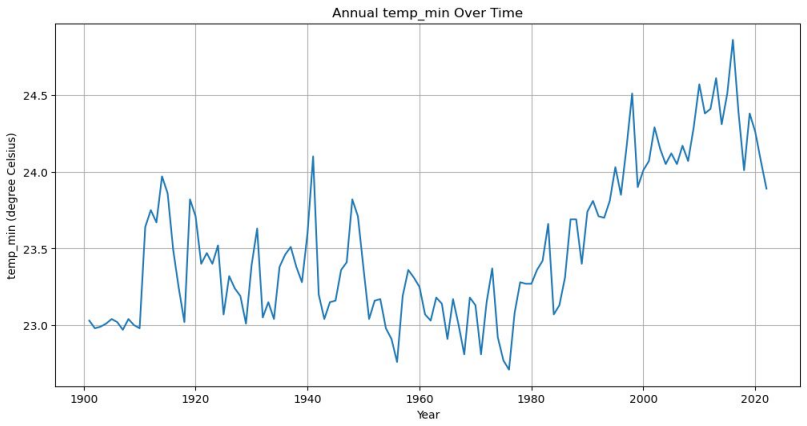




- Some outliers:
  - Year 1913 => High rainfall exceeding 3300mm
  - Year 1997 => Low rainfall of less than 1800mm
- Rainfall hovers in the range of 2200 to 2800mm

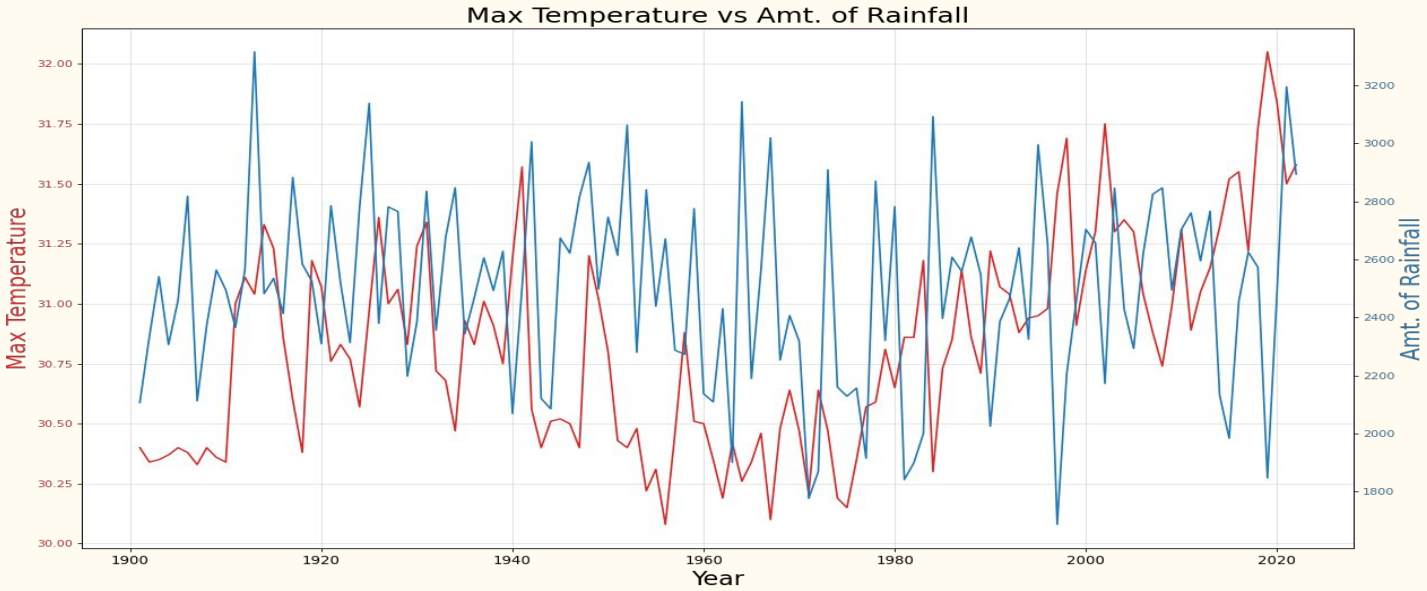
Data Exploration





- Range of the temperatures seems relatively big in these decades:
  - 1910-1920, 1940-1950, and 1970-1980
- But short range of temperatures in the experienced in 1900-1910 decade
- More outliers and higher temperatures experienced in the recent decades

Data Exploration



- Decreasing trend for both maximum temperature and rainfall in from year 1940 to 1960.
- Unusual years (e.g. 1934) which has high rainfall but low maximum temperature

Methodology

- Two analysis methods explored.

- First method-Machine Learning Model
  - Year as predictor and Climate Factors as response
  - Linear Regression Model
  - Support Vector Machine

- Second method-Time Series Analysis
  - ARIMA (Autoregressive Integrated Moving Average) model

Methodology

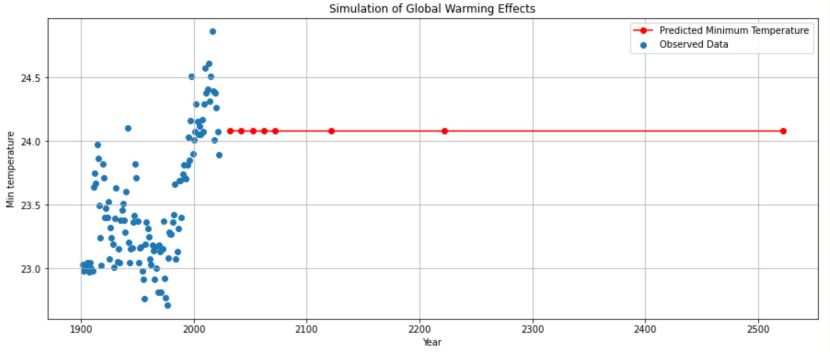
- For the machine learning models, the dataset is split into training and testing sets.
- Models are built using the train set and evaluated on the test set.
- The evaluation is done using MSE (Root Mean Squared Error), and R-Squared.
- The chosen model (linear regression or support vector machine) is used to forecast/predict the climate factor over the future years.

Methodology

- For the time series analysis we make use of the ARIMA model.
- Before implementing the model, initial analysis such testing whether the data is stationary or not, in addition initial model parameters were also determined by plotting the ACF and PACF plots.
- The parameters determined are,p, the number of autoregressive terms in the model, d,is the number of differentiations applied on the time series values and q, refers to the number of moving average terms in the model.
- The models implemented are evaluated using MSE (Mean Squared Error) and the residual plots of the climate factors.

Results-Forecasting

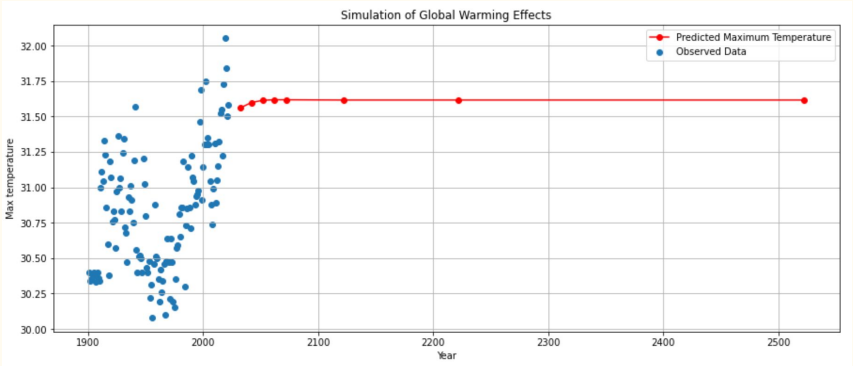
1. min_temp



ARIMA forecast

For the Time Series ARIMA model, we chose AutoRegressive Order=2, Differencing Order=1, and Moving Average Order=9, and the predicted future minimum temperature has a smooth trend and is expected to remain at 24.1 degrees Celsius for a long time.

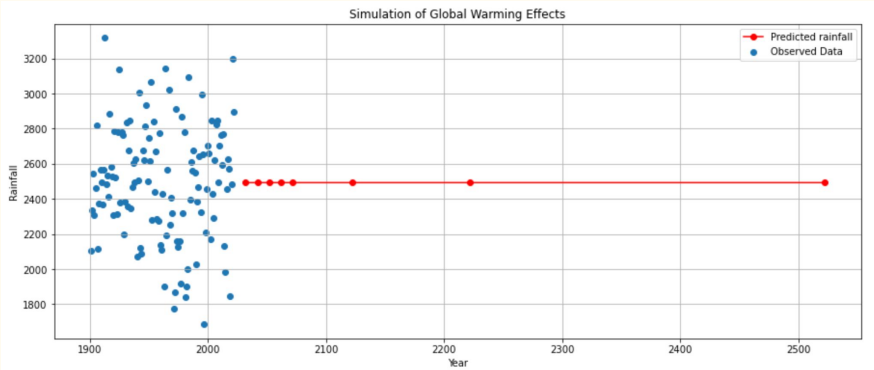
2. max_temp



ARIMA forecast

For the Time Series ARIMA model, we chose AutoRegressive Order=2, Differencing Order=1, and Moving Average Order=6, and the predicted future maximum temperature has a smooth trend and is expected to remain at 31.6 degrees Celsius for a long time.

3. rainfall



ARIMA forecast

For the Time Series ARIMA model, we chose AutoRegressive Order=1, Differencing Order=1, and Moving Average Order=2, and the predicted future total amount of rain collected for the year has a smooth trend and is expected to remain at 2500 mm just like the prediction of support vector regression.

Time series method-results

MSE:

models	min_temp	max_temp	rainfall
ARIMA regression	**0.051**	**0.070**	117640.904

ARIMA Parameters:

indicators	ACF/PACF	ARIMA parameters
min_temp	ACF: 9th-order trailing PACF: 2nd-order truncation	ARIMA (p=2, d=1, q=9)
max_temp	ACF: 6th-order trailing PACF: 2nd-order truncation	ARIMA (p=2, d=1, q=6)
rainfall	ACF: 1st-order truncation PACF: 1st-order truncation	ARIMA (p=1, d=1, q=2)

Machine learning method-results

Minimum temperature:

Models	R-squared	MSE
Linear regression	-5.943	0.440
Linear SVR	-5.627	0.420
Non-linear SVR (using RBF kernel)	-0.065	0.067

Maximum temperature:

Models	R-squared	MSE
Linear regression	-2.650	0.358
Linear SVR	-2.640	0.357
Non-linear SVR (using RBF kernel)	-2.481	0.342

Rainfall:

Models	R-squared	MSE
Linear regression	-0.102	140704.223
Linear SVR	-0.062	135581.193
Non-linear SVR (using RBF kernel)	-0.017	129825.557

Conclusion - short run effect

Analysis of 5-year average:

average minimum temperature projected to rise by 0.038 degree celsius

average maximum temperature projected to fall by 0.10 degree celsius

Significance:

→ decrease in maximum temperature could be **attributed to the La Niña effect** present in 2020-2022, which may lead to temporarily lower temperature.

→ Given the **high level of humidity in Singapore**, **temperature increase will be acutely felt.**

→ increase in temperature can **negatively impact cognitive performance**(Brink et al.,2020), which may affect productivity and economic growth.

Conclusion - short run effect

The average rainfall is projected to fall by 4% in 5 years

Significance:

→ risk to Singapore's already **limited water supply**

→ may pose risk of **inflationary pressures** in future

Conclusion - proposed action

To deal with temperature increase:

- Need for **stop-gap measures** to help residents deal with rising temperatures in their day-to-day activities.

  → there will likely be a need for greater heat insulation and flexibility in dressing in schools and workplaces.

To deal with decrease in rainfall:

- **diversify its water sources**, including **importing more from other countries.**

Overall, it is clear that **increased government expenditure** is required to help residents manage the effects of climate change in the short term.

Conclusion - long run effect

Average minimum temperature is projected to decrease by 0.20 degree celsius.

Average maximum temperature is projected to increase by 0.07 degree celsius.

average rainfall is projected to decrease by 0.12%.

Significance:

→ rise in temperature and **increased temperature volatility;**

→ temperature changes projected for next decade are more adverse than 5-year projections.

→ rainfall change projected for next decade is more moderate compared to the 5-year projection.

However, weather changes may become more pronounced with the increasing environmental damage brought on by expected rise in military efforts worldwide, such as in the Russia-Ukraine War and Israel-Palestinian conflict.

Conclusion - proposed actions

- **Stop-gap measures insufficient; systemic change required**

- Green Plan 2030: restore nature in its landscape

  → increased vegetation can help to reduce surface temperature and mitigate temperature changes in the next decade.

Future Works

- There is scope to **reflect greater complexity in how climate change data evolves over time** in machine learning models such as linear regression, SVR models.

  → engineered features such as moving average, **lagged predictors** in SVR, and the **El Niño and La Niña effects can be incorporated** into our models to improve predictions.

- Last but not least, there is merit in performing **cross-sectional multivariate analysis,**
  → e.g. by using measures of pollution as predictors to rainfall and maximum and minimum temperature to establish relationships between different predictors and the response variables.