Sydney Daily Rainfall

By: Muhammad Fadhil

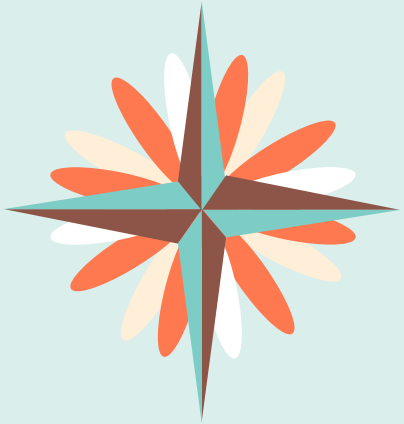
Table of Contents

Overview 01

EDA 02

03 Models

04 Conclusion



01

Overview

Problem Statement

Accurately predict the daily rainfall amount for the city of Sydney

Context

Human-caused climate change intensifies the heaviest downpours. Every 1°F rise also allows the atmosphere to hold 4% more water vapour.

As such the Department of Planning, Industry and Environment(DPIE) of New South Wales is particularly interested in Rainfall statistics in Sydney. Predicting the amount of Rainfall is imperative to the planning of the infrastructure of the city.
This would aid in the prevention of flooding and the accumulation of stagnant waters.

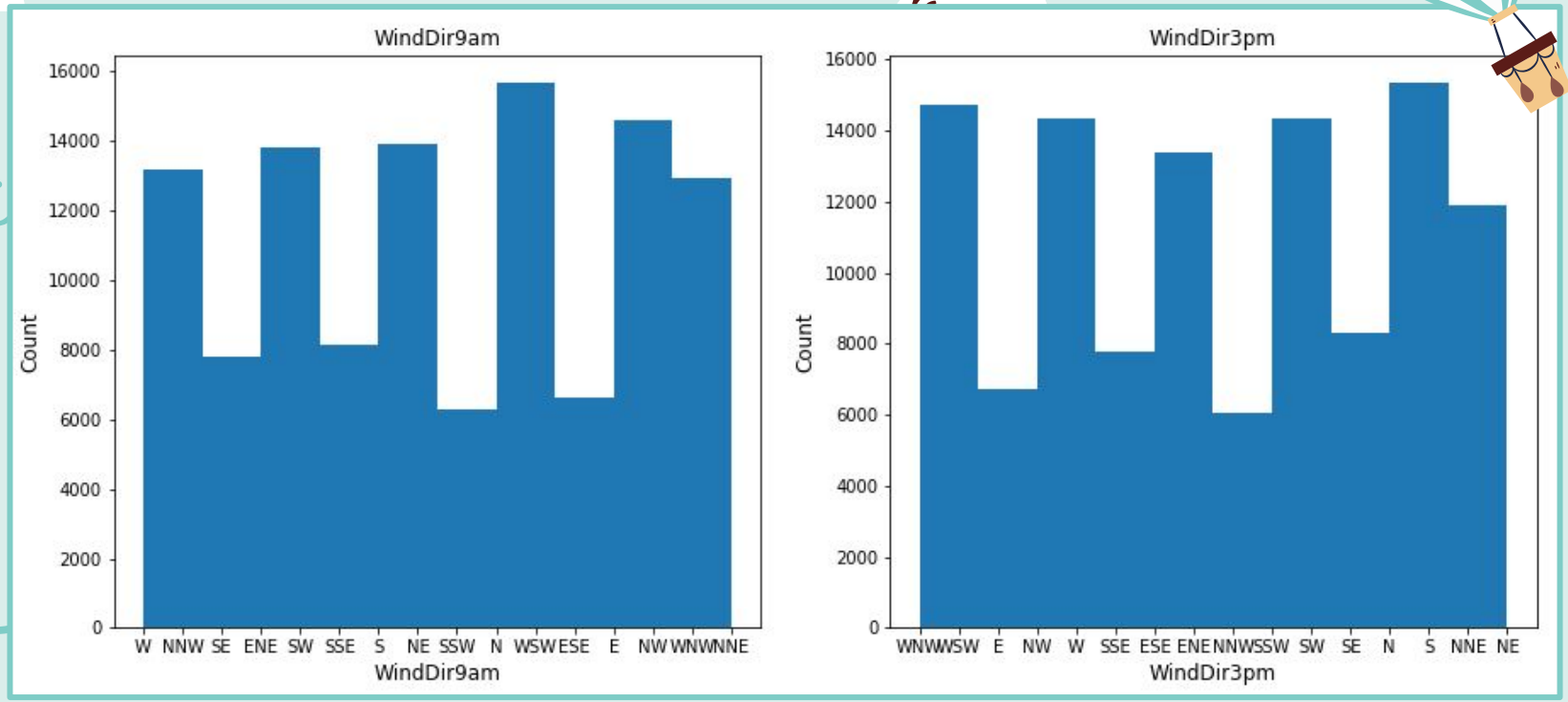
*dataset was obtained from kaggle and consisted of weather measurements for 49 cities in Australia from 2008 to 2017.

02

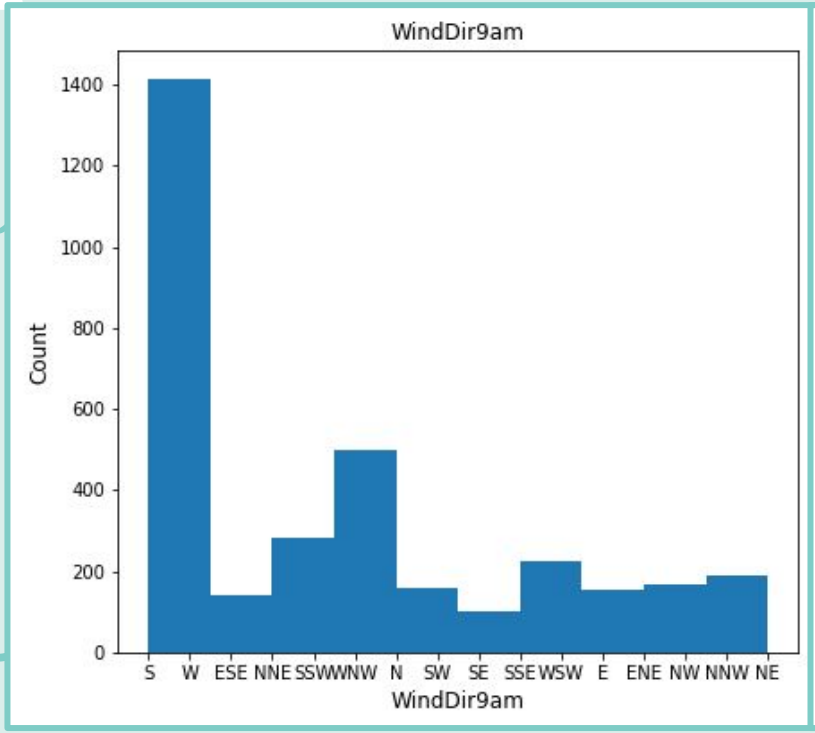
EDA

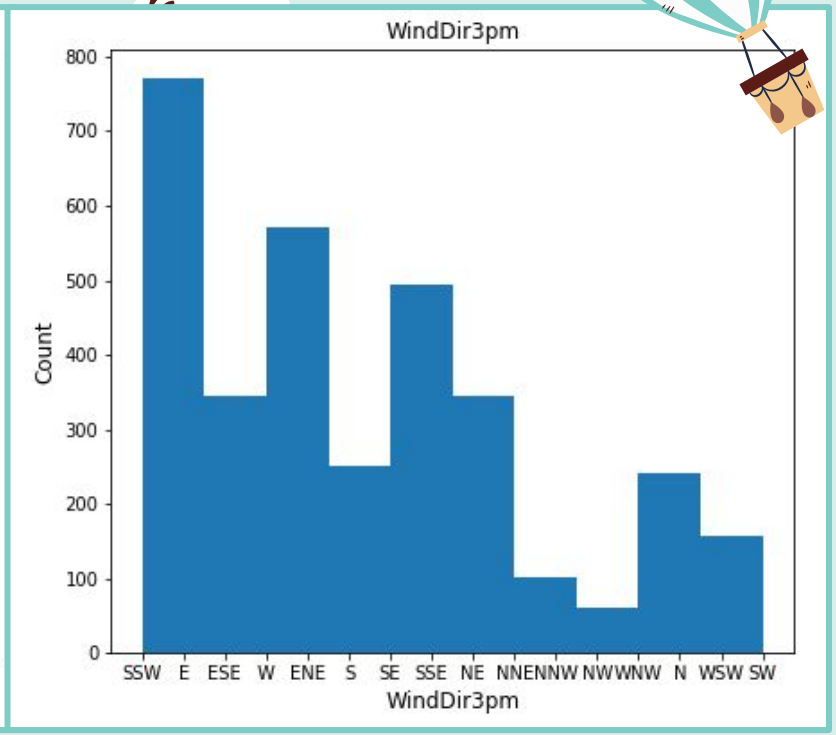
North America
Eurasia
Pacific ocean
Africa
Indian ocean
South America
Australia

Australian Winds



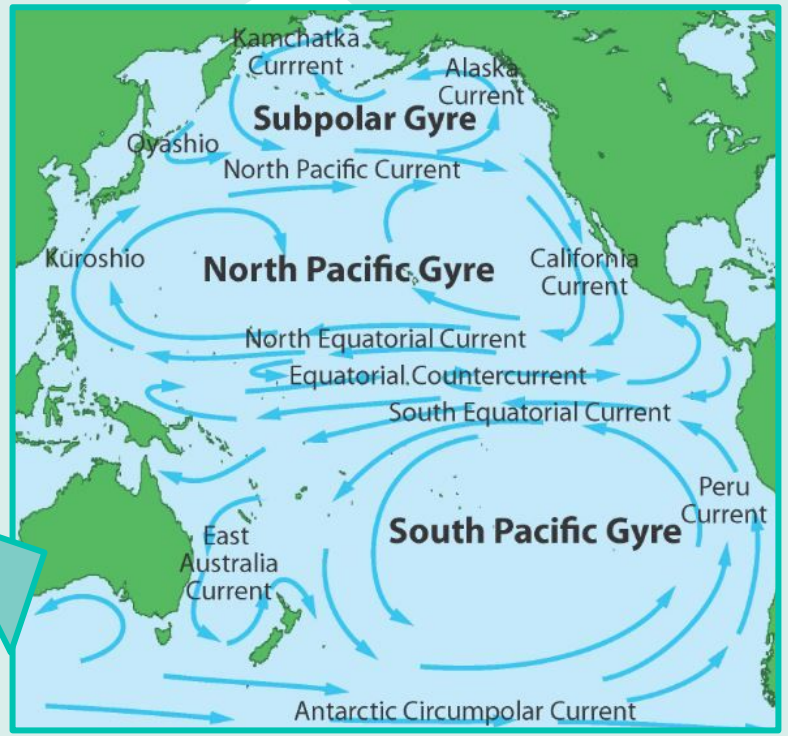
Sydney Winds



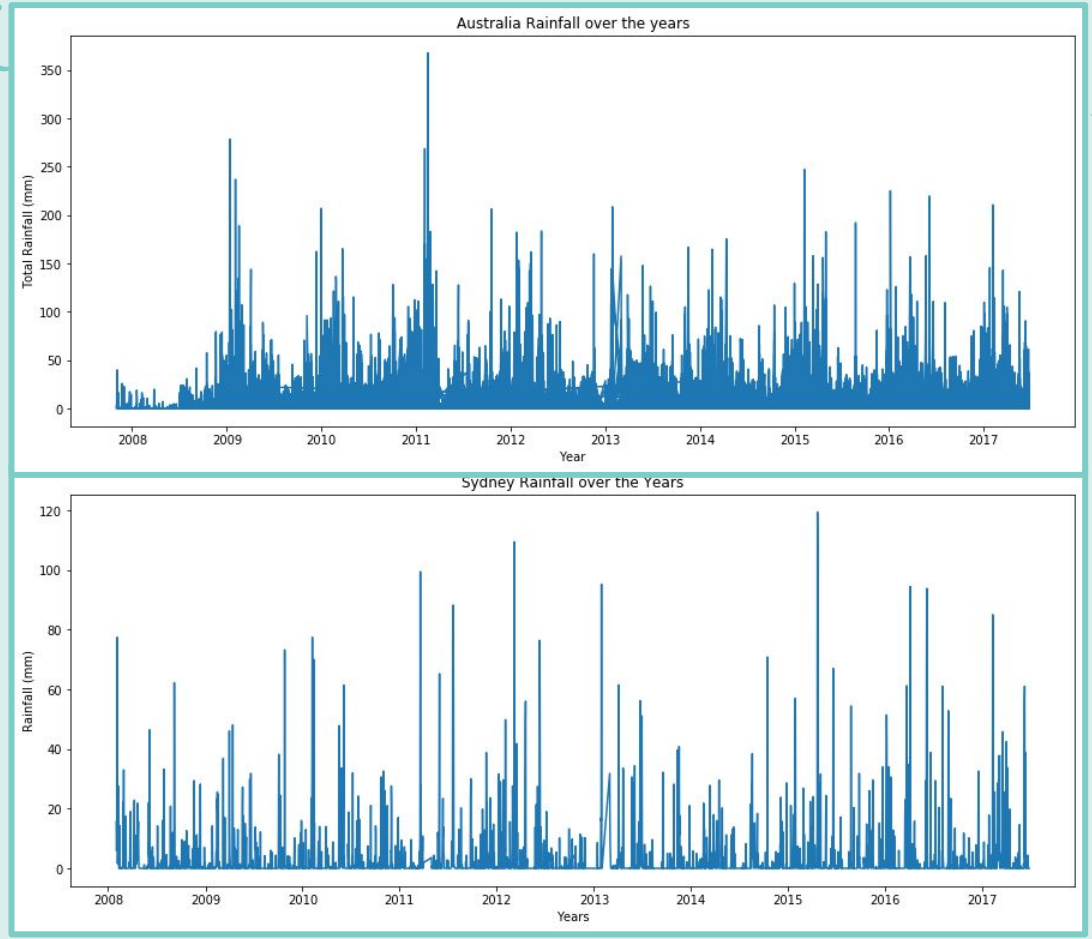


Natural Wind Currents

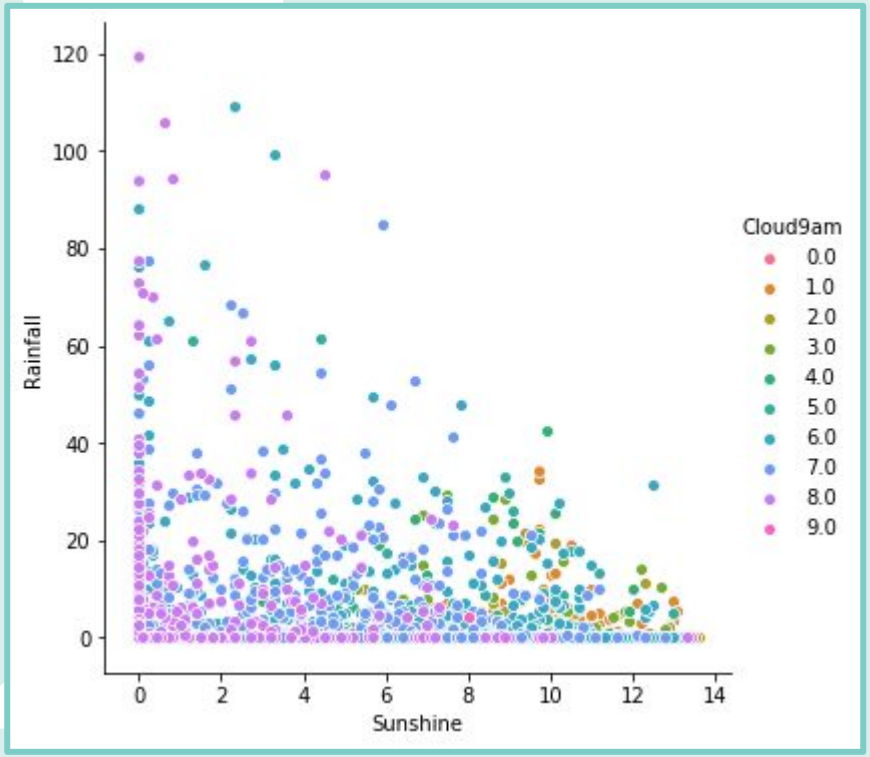
- Australia is located next to the Pacific Ocean
- Winds are influenced by natural currents
- Explains why there isn't much Northern Winds
- Even more so for Sydney



Australia vs Sydney Rainfall



Sydney's Cloud and Sunshine



- More Cloud Coverage means less Sunshine
- Sunshine inversely proportional to Rainfall
- Which makes sense intuitively, more cloud generally means higher chance or more rain

03

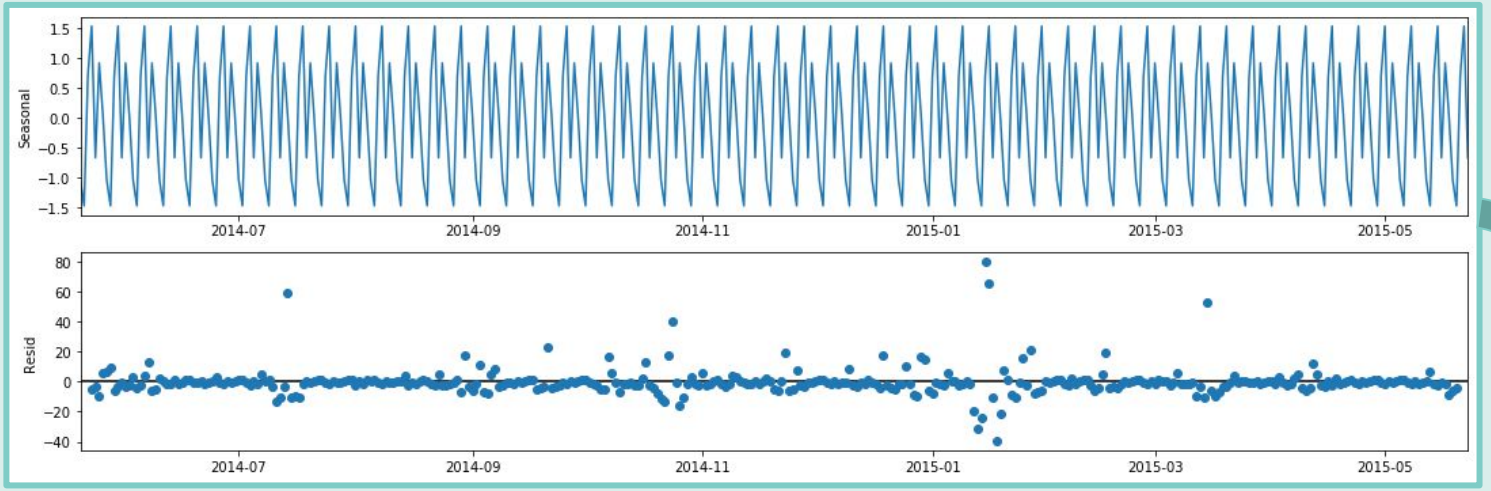
Modelling

3 Models

- SARIMAX
- Facebook’s Prophet
- Gated Recurrent Units

Seasonality

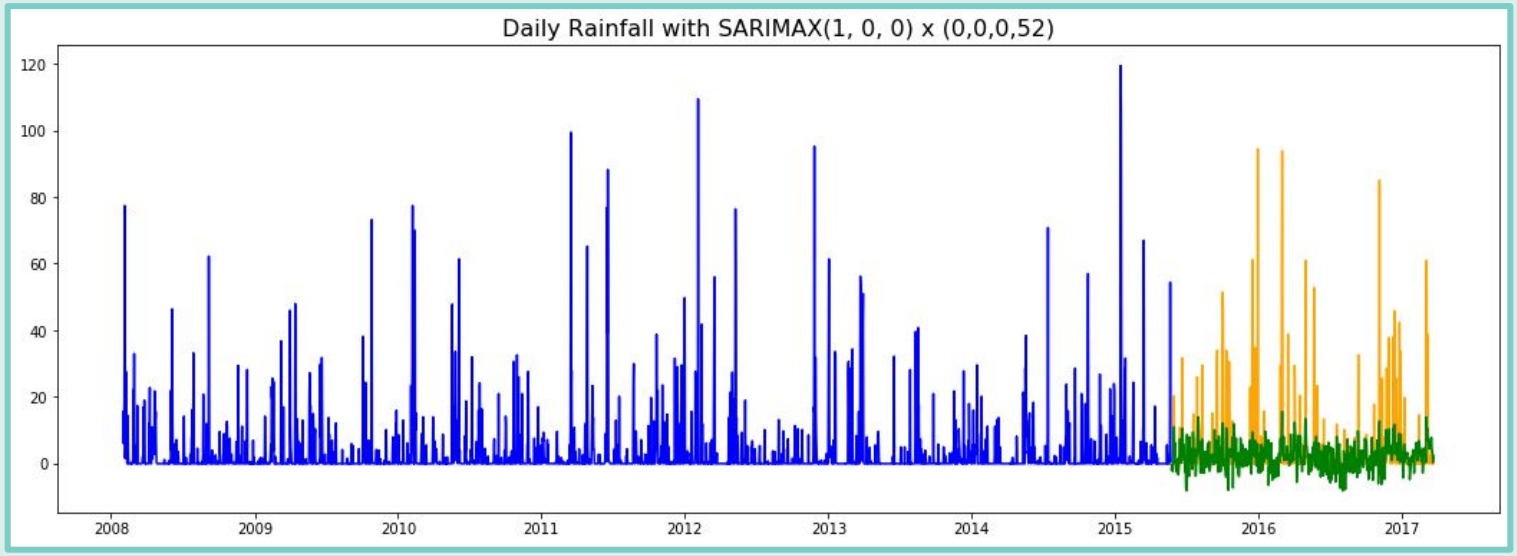
Statsmodel decompose



Weekly Seasonality

SARIMAX

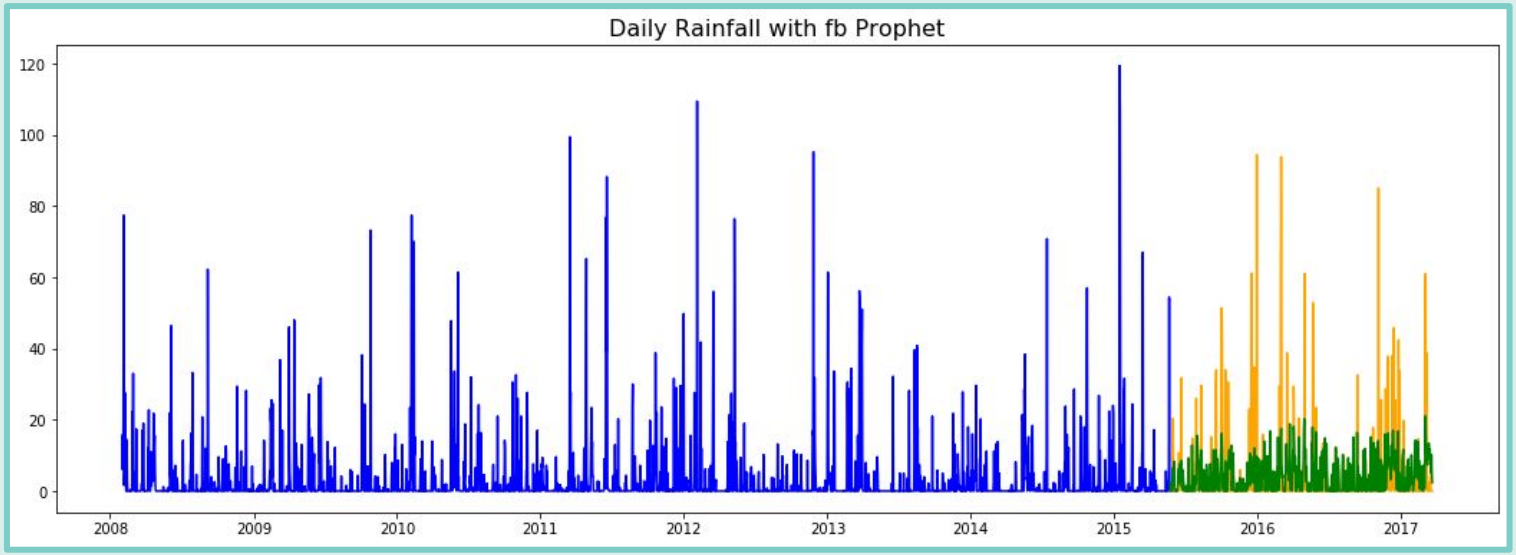
Baseline: 125.6



- SARIMAX did quite well with a Mean Squared Error of 99.1
- Model produced negative values which doesn't make sense for Rainfall
- Had to impute the seasonality inferred by our statsmodel

Fb's Prophet

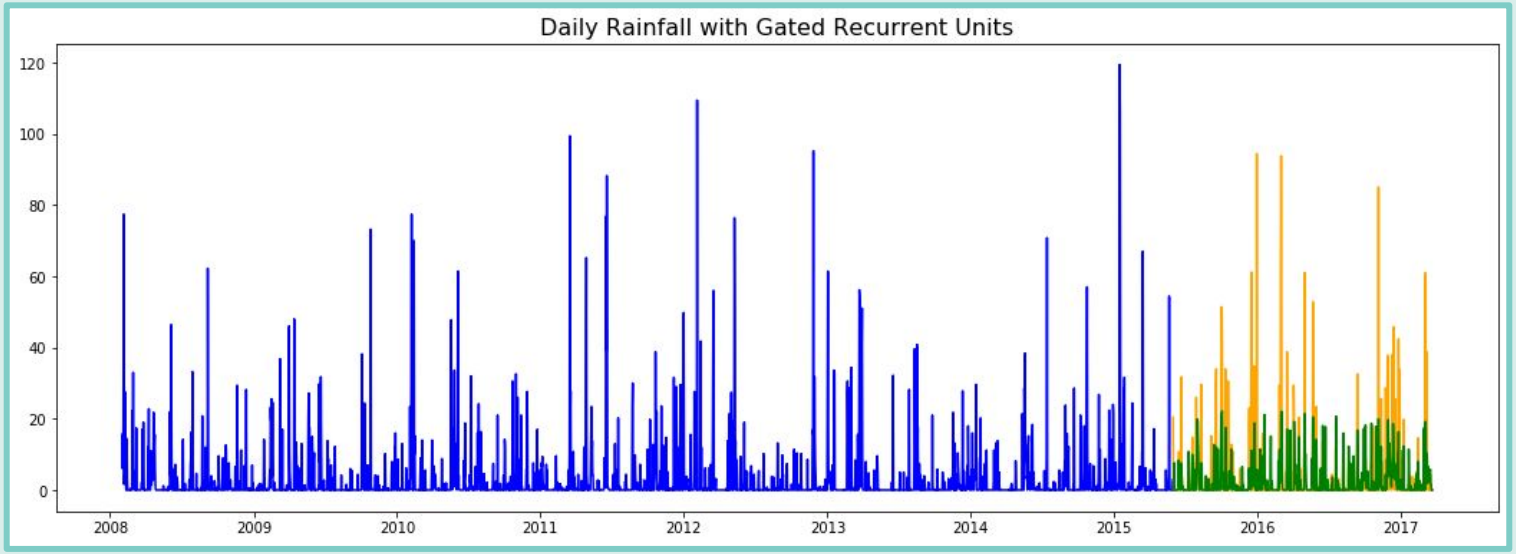
Baseline: 125.6



- Prophet did not do as well with a Mean Squared Error of 104.2
- Seasonality is inferred by Prophet and does not require imputation

Gated Recurrent Units

Baseline: 125.6



- GRU did a lot better with a Mean Squared Error of 96.4
- Seasonality is inferred by GRU and does not require imputation
- More Complex than our conventional time series

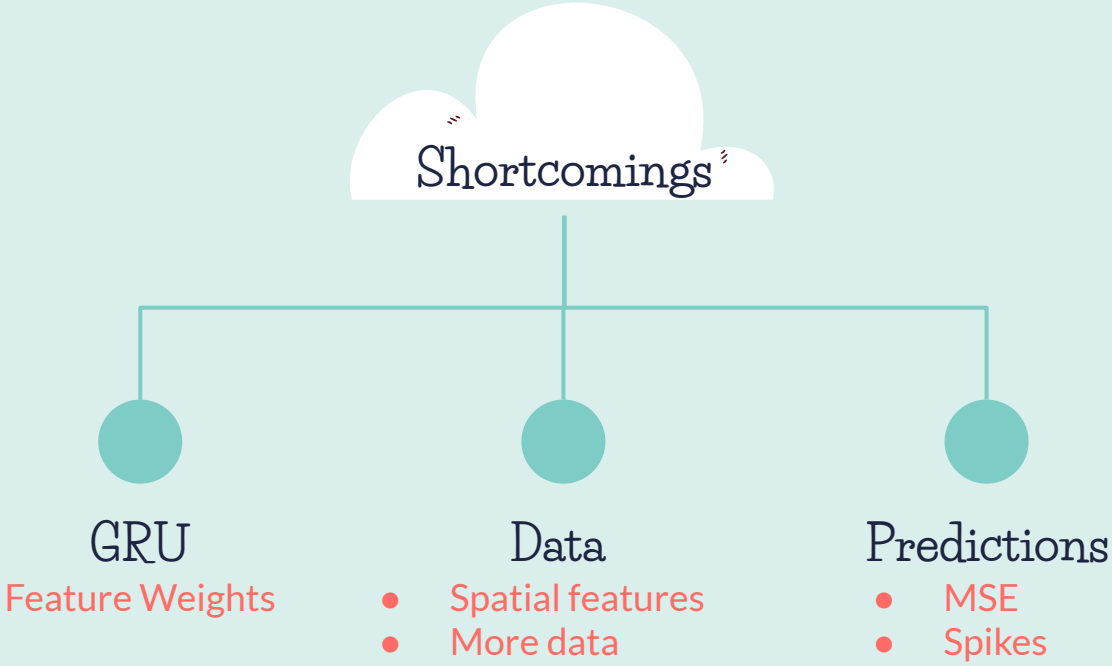
04

Conclusion

Summary

Method	Mean Squared Error
Baseline	125.6
SARIMAX	99.1
Prophet	104.2
GRU	96.4

Limitations



Thank You!

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

Please keep this slide for attribution