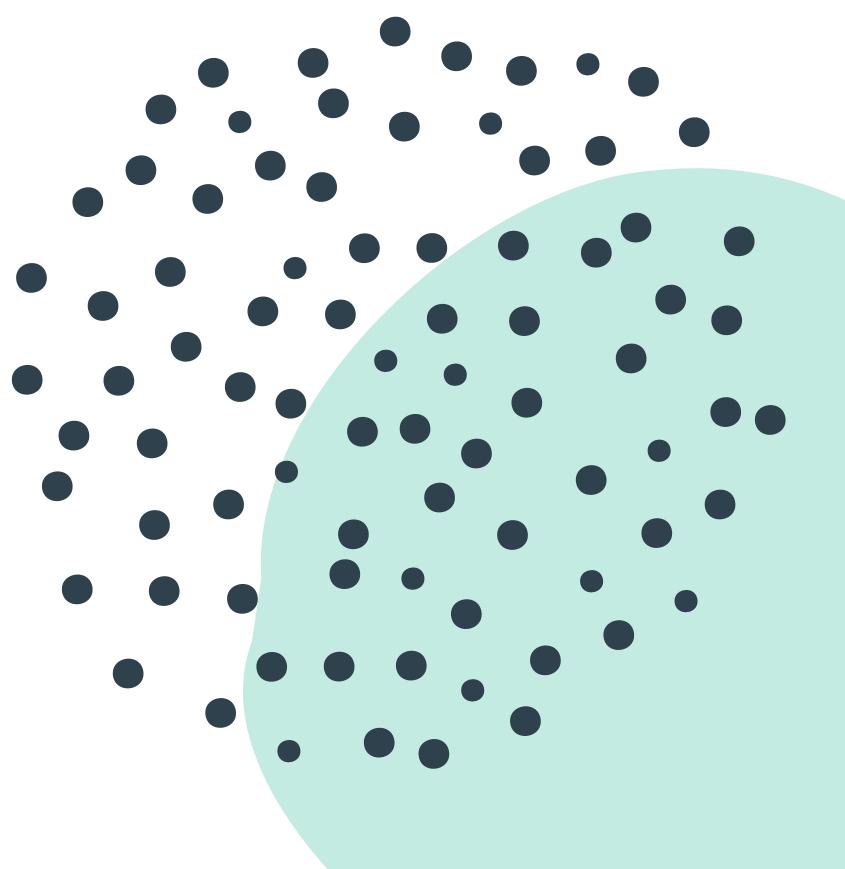


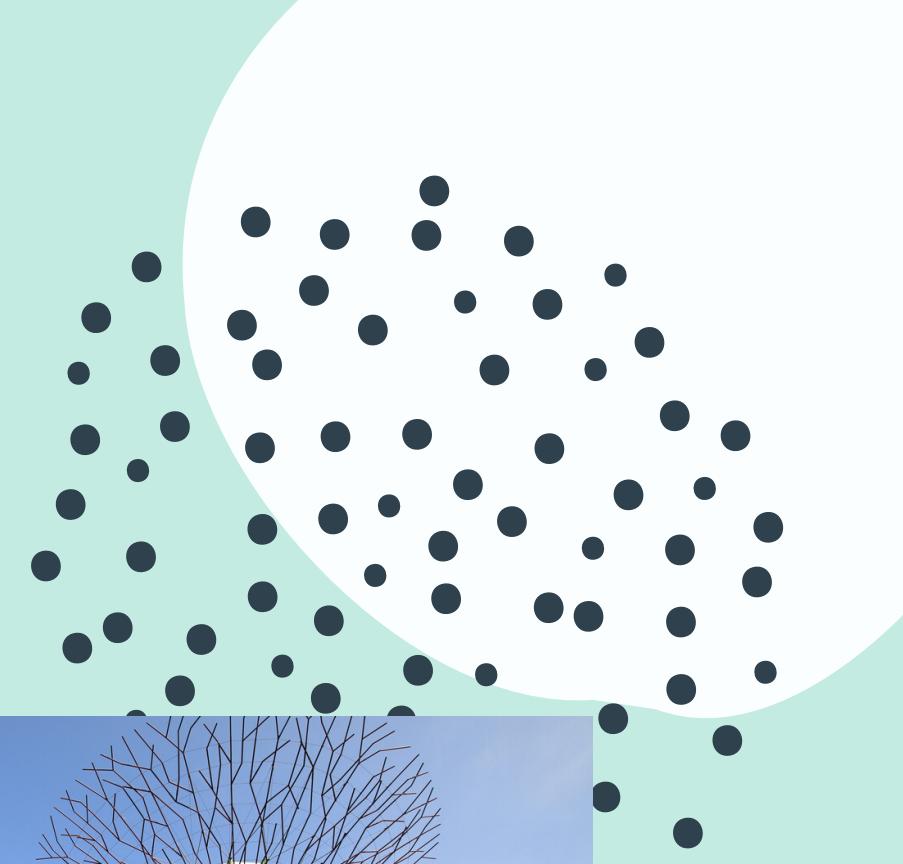
PREDICTING WEEKLY DENGUE CASES IN SINGAPORE

Using weather data and Google trends

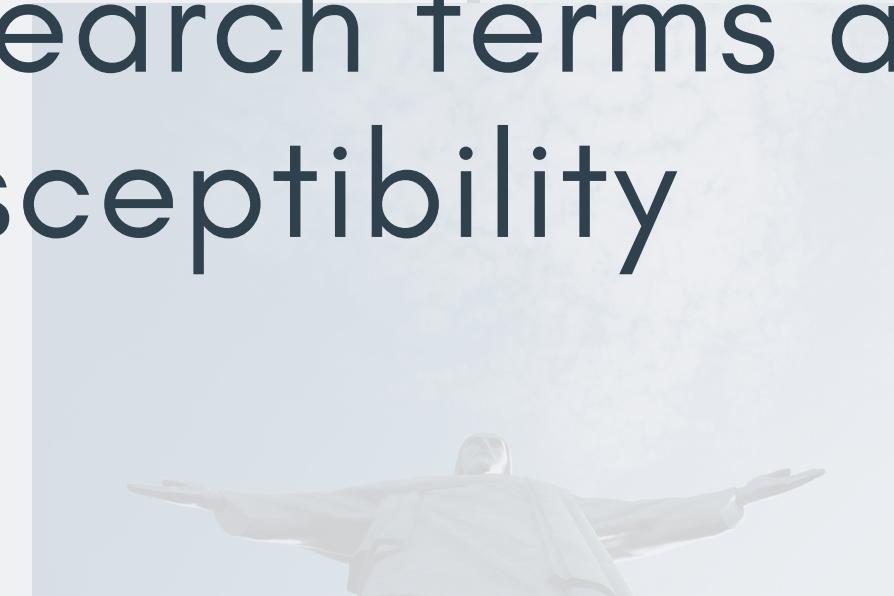
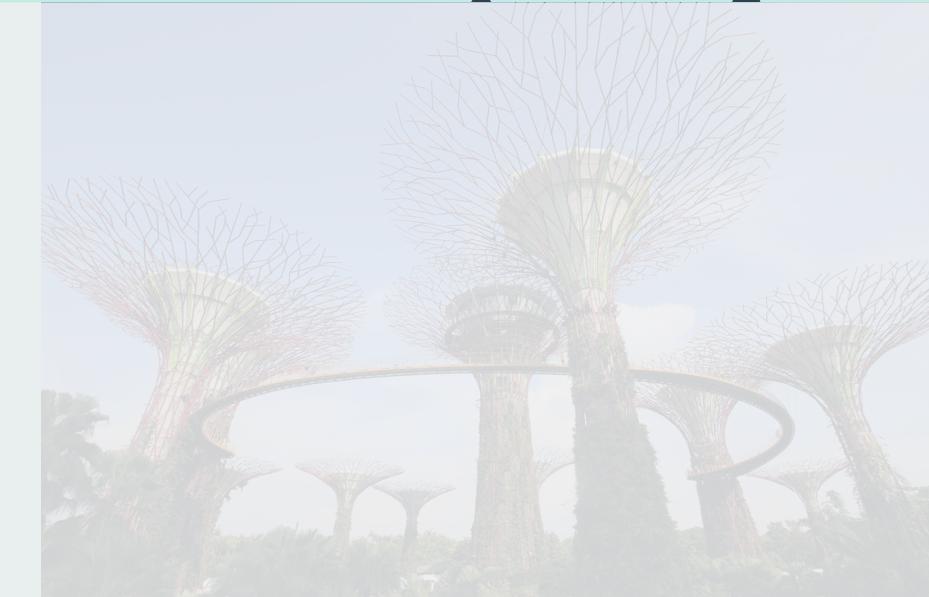
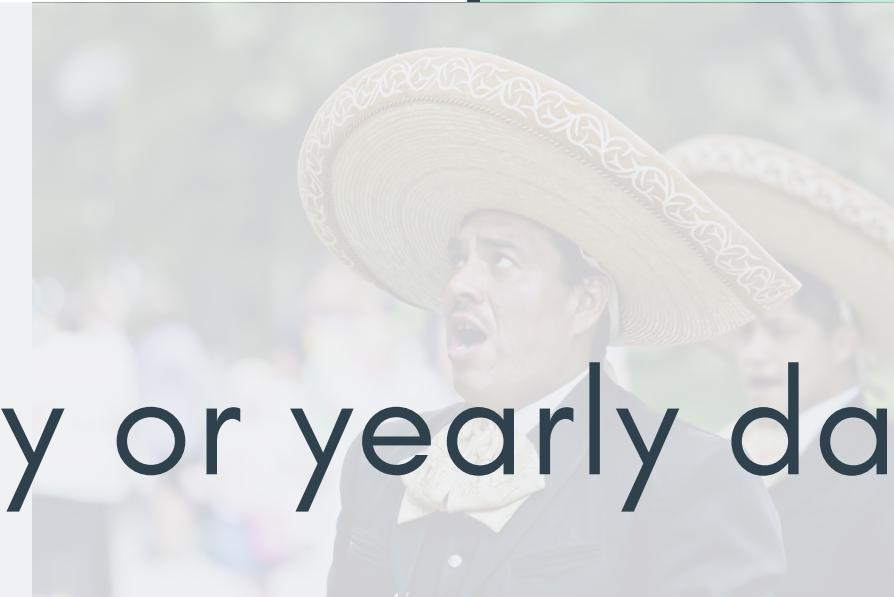




SEARCHING FOR DENGUE



SEARCHING FOR DENGUE



- Use monthly or yearly data
- 'Dengue' search terms are related to dengue susceptibility

TIME LAG DATA



REPORTED

Today

Case is reported
after you see a doc



INFECTED

1 week ago

- Fever
- Vomiting
- Nausea
- Headache
- Body aches
- Eye pain
- Rashes

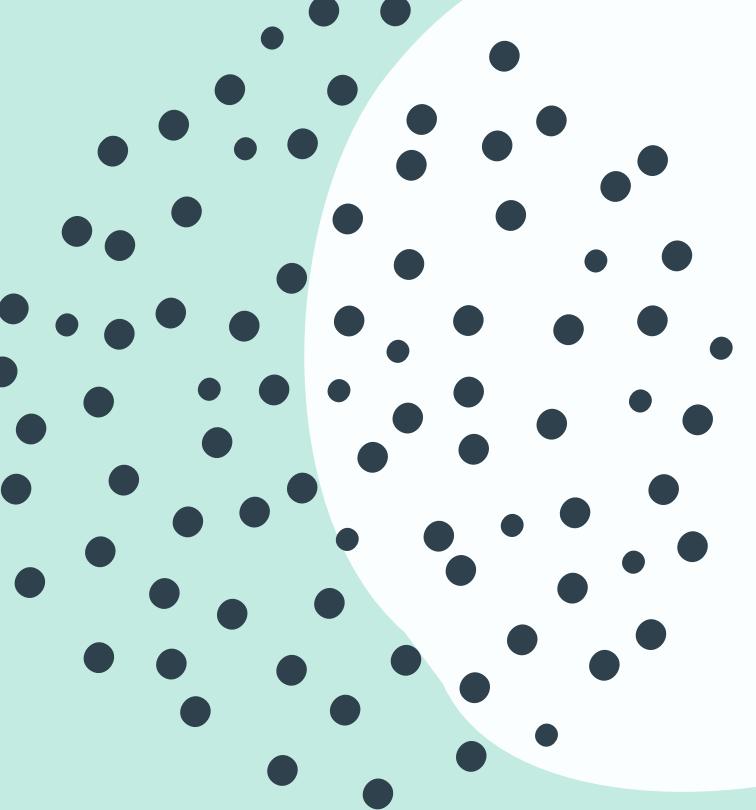


HATCHED

2 weeks ago

- Warm climate
- Stagnant water

METHODOLOGY



DATA

2014 - 2018
MOH weekly dengue
MSS daily weather data
Google Trends

MODEL

Simple linear regression

PACKAGES

Scikit-learn
StatsModel
BeautifulSoup
Selenium
EpiWeeks

MODEL

KEY TAKEAWAYS

1. Weekly dengue cases can be predicted using historic data.



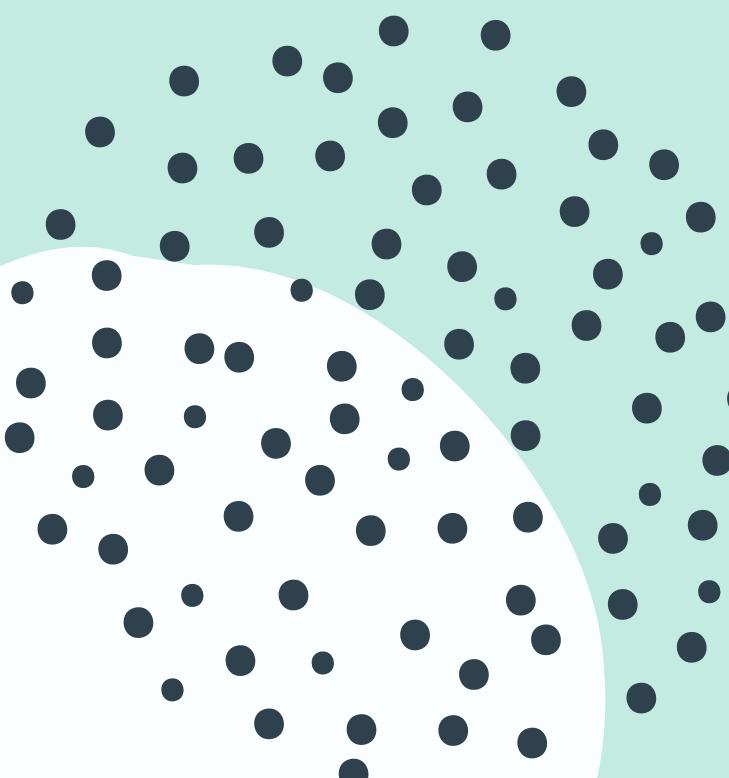
OLS Regression Results

Dep. Variable:	log_dengue_cases	R-squared:	0.782			
Model:	OLS	Adj. R-squared:	0.774			
Method:	Least Squares	F-statistic:	97.31			
Date:	Sat, 26 Oct 2019	Prob (F-statistic):	2.24e-59			
Time:	09:37:35	Log-Likelihood:	-98.646			
No. Observations:	198	AIC:	213.3			
Df Residuals:	190	BIC:	239.6			
Df Model:	7					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	4.8577	0.029	168.139	0.000	4.801	4.915
max_temp	0.0671	0.030	2.222	0.027	0.008	0.127
headache	-0.1550	0.032	-4.910	0.000	-0.217	-0.093
ache_pain	-0.1679	0.032	-5.194	0.000	-0.232	-0.104
eye_pain	0.0382	0.029	1.308	0.193	-0.019	0.096
rashes	0.0720	0.030	2.379	0.018	0.012	0.132
vomiting	-0.0405	0.031	-1.308	0.193	-0.102	0.021
log_dengue	0.5796	0.033	17.771	0.000	0.515	0.644
Omnibus:	1.102	Durbin-Watson:	2.185			
Prob(Omnibus):	0.576	Jarque-Bera (JB):	1.114			
Skew:	-0.178	Prob(JB):	0.573			
Kurtosis:	2.906	Cond. No.	1.82			

MODEL

KEY TAKEAWAYS

1. Weekly dengue cases can be predicted using historic data.
2. Increased search interest in dengue and rashes, as well as higher temperatures predicted increases in dengue cases.



OLS Regression Results

Dep. Variable:	log_dengue_cases	R-squared:	0.782			
Model:	OLS	Adj. R-squared:	0.774			
Method:	Least Squares	F-statistic:	97.31			
Date:	Sat, 26 Oct 2019	Prob (F-statistic):	2.24e-59			
Time:	09:37:35	Log-Likelihood:	-98.646			
No. Observations:	198	AIC:	213.3			
Df Residuals:	190	BIC:	239.6			
Df Model:	7					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	4.8577	0.029	168.139	0.000	4.801	4.915
max_temp	0.0671	0.030	2.222	0.027	0.008	0.127
headache	-0.1550	0.032	-4.910	0.000	-0.217	-0.093
ache_pain	-0.1679	0.032	-5.194	0.000	-0.232	-0.104
eye_pain	0.0382	0.029	1.308	0.193	-0.019	0.096
rashes	0.0720	0.030	2.379	0.018	0.012	0.132
vomiting	-0.0405	0.031	-1.308	0.193	-0.102	0.021
log_dengue	0.5796	0.033	17.771	0.000	0.515	0.644
Omnibus:	1.102	Durbin-Watson:	2.185			
Prob(Omnibus):	0.576	Jarque-Bera (JB):	1.114			
Skew:	-0.178	Prob(JB):	0.573			
Kurtosis:	2.906	Cond. No.	1.82			

MODEL

KEY TAKEAWAYS

1. Weekly dengue cases can be predicted using historic data.
2. Increased search interest in dengue and rashes, as well as higher temperatures predicted increases in dengue cases.
3. Decreased search interest in headaches, aches, pains were also predictive of increased dengue cases.



OLS Regression Results

Dep. Variable:	log_dengue_cases	R-squared:	0.782			
Model:	OLS	Adj. R-squared:	0.774			
Method:	Least Squares	F-statistic:	97.31			
Date:	Sat, 26 Oct 2019	Prob (F-statistic):	2.24e-59			
Time:	09:37:35	Log-Likelihood:	-98.646			
No. Observations:	198	AIC:	213.3			
Df Residuals:	190	BIC:	239.6			
Df Model:	7					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	4.8577	0.029	168.139	0.000	4.801	4.915
max_temp	0.0671	0.030	2.222	0.027	0.008	0.127
headache	-0.1550	0.032	-4.910	0.000	-0.217	-0.093
ache_pain	-0.1679	0.032	-5.194	0.000	-0.232	-0.104
eye_pain	0.0382	0.029	1.308	0.193	-0.019	0.096
rashes	0.0720	0.030	2.379	0.018	0.012	0.132
vomiting	-0.0405	0.031	-1.308	0.193	-0.102	0.021
log_dengue	0.5796	0.033	17.771	0.000	0.515	0.644
Omnibus:	1.102	Durbin-Watson:	2.185			
Prob(Omnibus):	0.576	Jarque-Bera (JB):	1.114			
Skew:	-0.178	Prob(JB):	0.573			
Kurtosis:	2.906	Cond. No.	1.82			

MODEL

KEY TAKEAWAYS

1. Weekly dengue cases can be predicted using historic data.
2. Increased search interest in dengue and rashes, as well as higher temperatures predicted increases in dengue cases.
3. Decreased search interest in headaches, aches, pains were also predictive of increased dengue cases.
4. Model can be generalised. Adj. R-squared for test: **0.760**



OLS Regression Results

Dep. Variable:	log_dengue_cases	R-squared:	0.782			
Model:	OLS	Adj. R-squared:	0.774			
Method:	Least Squares	F-statistic:	97.31			
Date:	Sat, 26 Oct 2019	Prob (F-statistic):	2.24e-59			
Time:	09:37:35	Log-Likelihood:	-98.646			
No. Observations:	198	AIC:	213.3			
Df Residuals:	190	BIC:	239.6			
Df Model:	7					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	4.8577	0.029	168.139	0.000	4.801	4.915
max_temp	0.0671	0.030	2.222	0.027	0.008	0.127
headache	-0.1550	0.032	-4.910	0.000	-0.217	-0.093
ache_pain	-0.1679	0.032	-5.194	0.000	-0.232	-0.104
eye_pain	0.0382	0.029	1.308	0.193	-0.019	0.096
rashes	0.0720	0.030	2.379	0.018	0.012	0.132
vomiting	-0.0405	0.031	-1.308	0.193	-0.102	0.021
log_dengue	0.5796	0.033	17.771	0.000	0.515	0.644
Omnibus:	1.102	Durbin-Watson:	2.185			
Prob(Omnibus):	0.576	Jarque-Bera (JB):	1.114			
Skew:	-0.178	Prob(JB):	0.573			
Kurtosis:	2.906	Cond. No.	1.82			

CONCLUSION

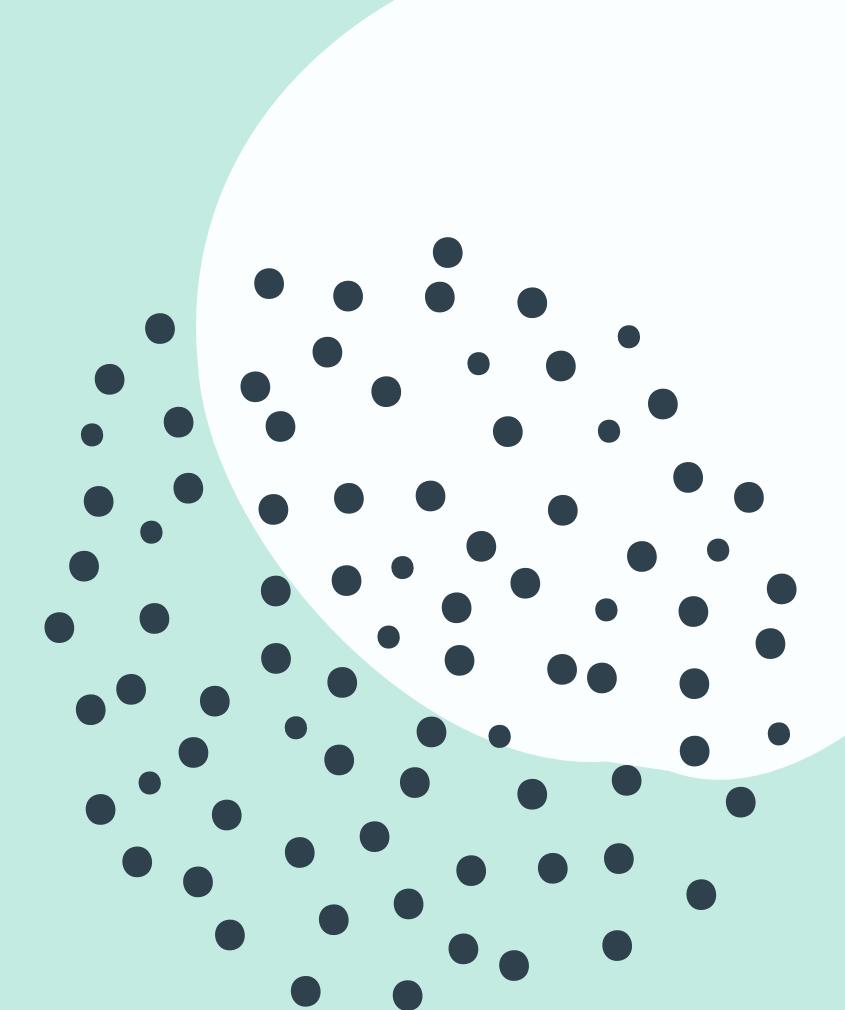
There is potential in predicting dengue cases using past data.

'Dengue' searches was the strongest predictors of dengue cases

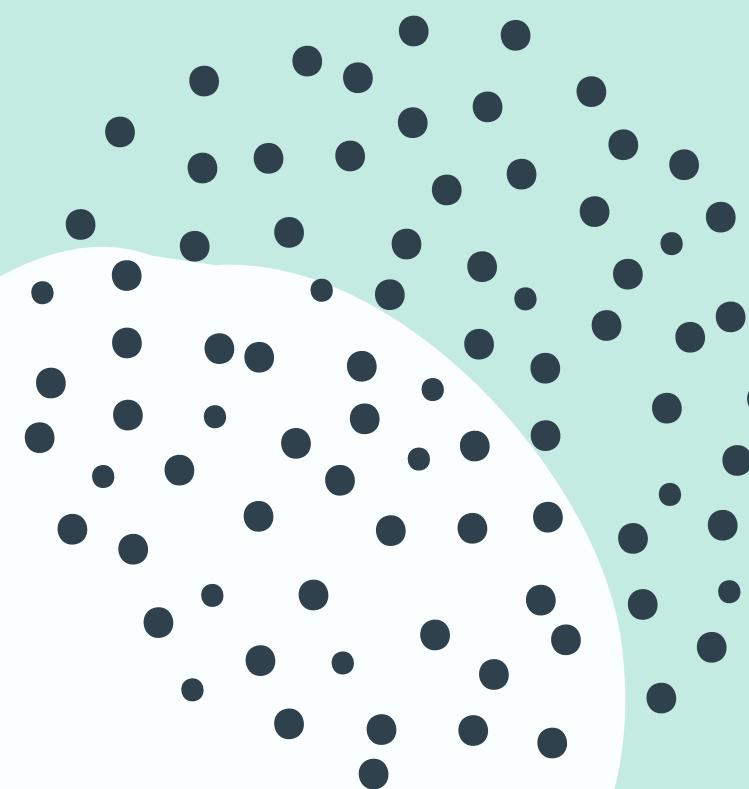
Search interest could be fuelled by media reports on dengue. Thus, searches of 'dengue' could be positively related to increases in dengue breeding clusters or cases.

Negative coefficients of dengue-related symptoms

Awareness of dengue is high in Singapore. Hence, people could be searching for 'dengue' instead of its symptoms when flu-like symptoms follow a mosquito bite, or occur during dengue outbreaks.



FUTURE WORK



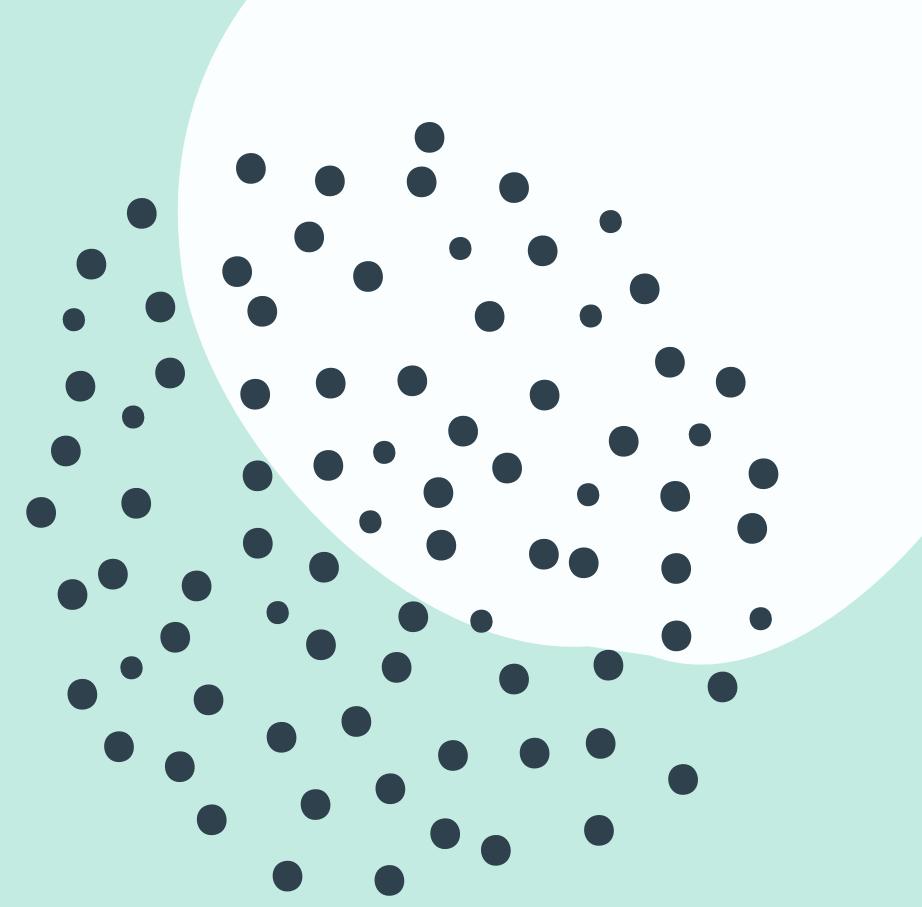
UNDERSTANDING SEARCH PATTERNS

Do dengue awareness levels
really affect search patterns?

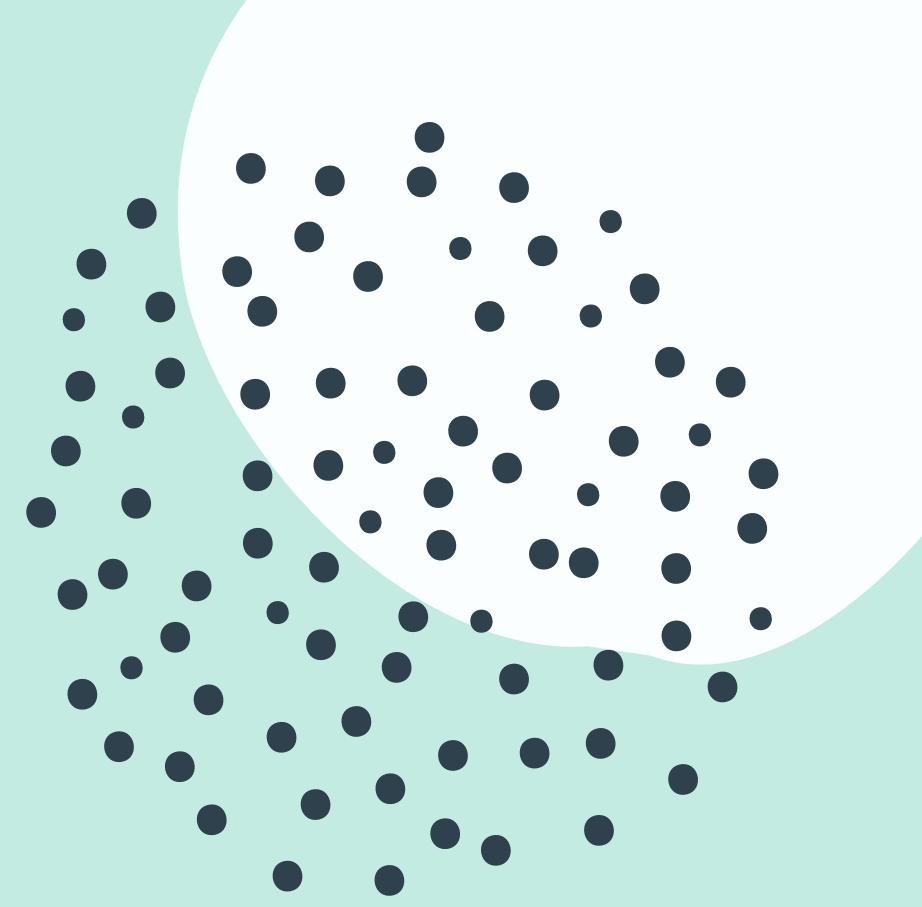
GET MORE DATA

Enable more flexible time lags
(e.g. rolling 5-day lead time)

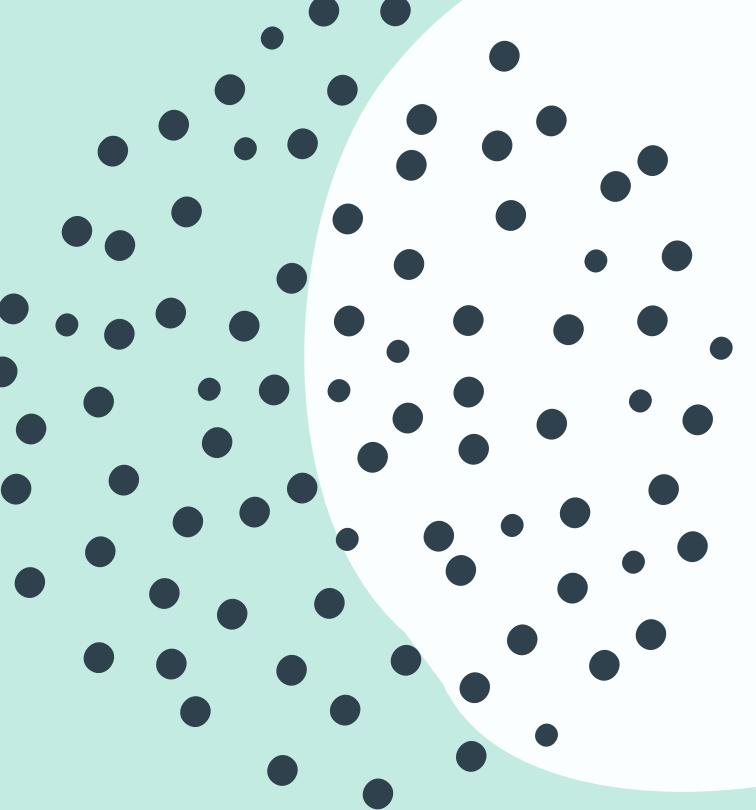
THANK YOU



APPENDIX



DATA SOURCES



WEEKLY DENGUE CASES

<https://data.gov.sg/datasets/weekly-number-of-dengue-and-dengue-haemorrhagic-fever-cases>

DAILY WEATHER DATA

<http://www.weather.gov.sg/climate-historical-daily/>

WEEKLY SEARCH TERMS

<https://trends.google.com/trends/trendingsearches/daily?geo=SG>

WEATHER FEATURES

Max Temperature
Min Temperature
Mean Temperature
Rainfall

SEARCH TERM FEATURES

Dengue	Fever
Ache/pain	Nausea
Eye pain	Vomiting
Headache	

CHECKING MODEL ASSUMPTIONS

1. Error terms have a constant variance
2. Errors are normally distributed around mean = 0
3. Errors are uncorrelated across observations

