S.T. Yau High School Science Award (Asia)
2020

Research Report

The Team

Registration Number: Econ - 030

Name of team member: Ayush Goel
School: The International School, Bangalore
City, Country: Bangalore, India

Name of team member: Sanjana Ganesh
School: Neev Academy, Bangalore
City, Country: Bangalore, India

Name of supervising teacher: Miss Sumita Tambi
Position: Teacher, Economics Department
School/Institution: The International School, Bangalore
City, Country: Bangalore, India

Title of Research Report

A Novel Near Real-time System to Estimate Impact of Sudden Economic Shocks (such as COVID-19) on the Economy

Date

Aug 31st, 2020

A Novel Near Real-time System to Estimate Impact of Sudden Economic Shocks (such as COVID-19) on the Economy

Ayush Goel, Sanjana Ganesh

Abstract

As the novel coronavirus progressed from Wuhan to various parts of the globe, governments were faced with a tradeoff: to go into lockdown to save lives but hurt the economy, or stay open to save the economy but lose lives. At the time, those in favour of the lockdown had empirical data to show the lethality of the virus and the need to go into lockdown. Whereas those against lockdowns didn't have data to show the severe reduction in economic activity that would follow the lockdown. Hence, numerous governments enforced lockdowns and realized the need to re-open much after the damage had already been done. This delayed realization was purely due to the lack of a real-time system to quantitatively gauge the economic impact of various policies. In this paper, we develop a near real-time framework to estimate the economic impact of such sudden shocks at any geographical level (city, state or country). We use novel data sources, specifically satellite recorded Night Light and Electricity Consumption data. Such alternate sources are advantageous as they can answer economic questions related to isolated locations and the informal economy (a sizable constituent of developing economies). We demonstrate the applicability of our framework across a variety of cities across countries (both developing and developed) and present our results specifically for the COVID-19 pandemic. We quantitatively estimate the economic decline in various cities, how government policies affected this decline, and the impact on intercity transport. Using Night Light, we estimated a GDP decline of 12.3% for New York and Miami (Apr - June quarter), and the U.S. government reported 7.2% at the country level. This difference is expected since these would have had a higher decline than the overall country. Moreover, our framework could make these predictions one month before the official numbers were released, reflecting the framework's accuracy and real-time nature. This framework could revolutionize the way governments, policymakers and stakeholders receive economic data, leading to more rational and proactive decision making.

Keywords: *Night Light, Real-Time analysis, Economic Impacts, Time Series Analysis, COVID-19, Satellite data, Alternate data sources, Electricity consumption*

Commitments on Academic Honesty and Integrity

We hereby declare that we

1. are fully committed to the principle of honesty, integrity and fair play throughout the competition.
2. actually perform the research work ourselves and thus truly understand the content of the work.
3. observe the common standard of academic integrity adopted by most journals and degree theses.
4. have declared all the assistance and contribution we have received from any personnel, agency, institution, etc. for the research work.
5. undertake to avoid getting in touch with assessment panel members in a way that may lead to direct or indirect conflict of interest.
6. undertake to avoid any interaction with assessment panel members that would undermine the neutrality of the panel member and fairness of the assessment process.
7. observe all rules and regulations of the competition.
8. agree that the decision of YHSA(Asia) is final in all matters related to the competition.

We understand and agree that failure to honour the above commitments may lead to disqualification from the competition and/or removal of reward, if applicable; that any unethical deeds, if found, will be disclosed to the school principal of team member(s) and relevant parties if deemed necessary; and that the decision of YHSA(Asia) is final and no appeal will be accepted.

(Signatures of full team below)

ayush

Name of team member: Ayush Goel

Sanjana

Name of team member: Sanjana Ganesh

Name of supervising teacher: Sumita Tambi

Noted and endorsed by (signature) Name of school principal: Dr. Caroline Pascoe, The International School Bangalore	Noted and endorsed by (signature) Name of school principal: Kavita Gupta Sabharwal, Neev Academy

Table of Contents

1. Introduction

1.1 Context

Policy Makers and Economists have to actively track the state of the global economy, especially during times of crises, such as natural disasters, financial crashes and economic recessions. Often, these crises prevent GDP data from being collected in real-time, especially in the case of natural disasters. However, it is during these times that this data is crucial and needed during the decision making process for governments, policymakers and economists.

Moreover, governments often report inaccurate figures for financial indicators, either due to their lack of infrastructure to support a comprehensive data collection or to portray the economy as superior to what it is. Here, economists and stakeholders need an alternate transparent perspective to understand the financial situation of the country as the quality of official data is not satisfactory. Moreover, this alternate perspective should also be frequently accessible, at least monthly, for it to be usable for various financial decisions.

Economists and policymakers are then forced to make tradeoffs, without access to important accurate GDP data, which could have extreme consequences. In such scenarios, a wrong decision is often made purely due to the lack of real-time or quality data. This deficiency is only apparent in hindsight.

The current system of measuring GDP, therefore, is not adequate on its own during periods of economic shock, due to the lack of transparency and non-real-time nature. Hence, the need arises for a new way to estimate GDP changes during these times by using alternate data sources that are highly correlated with GDP.

1.2 Problem Statement

We will be using data from two different alternate data sources to estimate GDP data in real-time: Night Light Radiance from the VIIRS satellite, and Electricity Consumption Data from the official data sources of respective countries. Both data sources are measured and updated regularly (Night Light daily, and Electricity Consumption monthly). We then develop a statistical framework for cleaning, processing, and analyzing this recorded data to make accurate predictions for GDP changes that can be calculated multiple times within a quarter, making it a near real-time system. Finally, we demonstrate the wide applicability of our framework by analyzing the impact of COVID on various sized cities and states, from both developing and developed countries, and how different government responses to the pandemic resulted in different levels of negative impact on the economy. Such information could help policymakers make proactive decisions in times where the situation evolves rapidly, and quarterly results are too delayed to make use of.

2. Related Work

2.1 Night Light

The seminal paper by Henderson et al. (American Economic Review 2012) was one of the first studies that showed a strong correlation between Night Light and GDP. Using data from the old DMSP satellite (Defense Meteorological Satellite Program), they were able to develop an end to end statistical framework to estimate GDP for countries with no or low-quality economic data. However, the data used was from the old DMSP satellite, which had a saturation value of 63, and hence couldn't differentiate between locations beyond a brightness of 63.

Further research by Pinkovskiy et al. (2015) focused on using night light as a means to compare the accuracy of GDP data across countries. They built on the idea that night light is a good proxy for GDP, and that it is easily comparable across countries. They then evaluated results from government reports with their results, derived using night light, and tried to gauge the quality of GDP data.

Also taking advantage of the ability to capture regions of different sizes (country, state, city) using night light, further research (Prakash et al. RBI Occasional Papers) showed that night light can also be used to accurately understand the economy at any scale, from villages, towns and cities to states and countries, with special focus on India.

However, research doesn't yet extrapolate such correlations between night light and GDP to analyze the impact of sudden economic shocks, where timely economic data (crucial to decision making) is not available, which is a major contribution of this paper.

In parallel, researchers have also worked on improving night light data quality from raw sensor data by removing unwanted effects from natural phenomena (Elvidge, et al.) such as moonlight reflectance, cloud scattering, background noise, and other outlier measurements.

However, as we show in this paper, when using night light for a time series analysis of the economy, there is an additional data processing phase required when analyzing the data.

2.2 Electricity Consumption

The seminal paper by Ferguson, Wilkinson and Hill (2000) was one of the first studies to show a strong correlation between electricity consumption (measured as kWh) and GDP. Using the data of over 100 countries, developing and developed, the study encompassed 99% of the current global economy. The study looked at each country's electricity consumption and GDP, to show that all countries showed a positive correlation between electricity consumption and GDP, with developed nations having a stronger correlation. Moreover, looking at the entire global economy, they showed a stronger correlation between electricity use and GDP. From this paper, the authors concluded that in the technological world of the 20th century, economic growth comes concurrently with electricity consumption.

Building on the works of this paper, a paper by Christopher M. Chima and Rodney Freed (International Business & Economics Research Journal 2005) showed the strong correlation between electricity consumption and GDP in the USA from 1949 – 2003, as well as how electricity consumption is a required input for, and a function of, economic growth.

A paper published by the Latin American Journal of Economics in 2013, looked at 10 Latin American countries and proved the strong positive correlation between electricity consumption and economic growth for all the sample countries. Another paper by Asit Mohanty (International Journal of Economics and Finance, 2015) showed how for India, electricity consumption fuels economic growth in both the short and long run, seen through a time series analysis. A paper by Enisa Džananović and Sabina Dacic Lepara published in 2017 looked at the relationship between electricity consumption and GDP in Southeast European Countries while showing a strong, significant positive correlation between GDP per capita and electricity consumption. Several other papers focused on demonstrating the strong correlation between electricity consumption and GDP for countries around the world.

While several papers (like the ones above) have shown the strong correlation between electricity consumption and GDP, research hasn't yet focused on using empirical factors like electricity consumption to study the impact of sudden economic shocks, periods when near real-time economic data (crucial to decision making) is not available.

3. Data Sources

3.1 Night Light

Global Night Light data was originally collected by the Defense Meteorological Satellite Program (DMSP), which used satellites to collect various weather indicators. One of these sensors had the capability of measuring light intensity. When measurements were taken at night, light emissions by cities, villages, etc were clearly visible. However, these satellites were later discontinued, but all recorded data is still publicly available (from 1992 to 2013).

These satellites have been replaced with the Suomi National Polar Partnership (SNPP) satellite, run by NASA. Data collected by the onboard Visible Infrared Imaging Radiometer Suite (VIIRS) sensor is also able to capture light emissions apart from many other indicators. Data collected by this sensor is superior in quality as compared to the DMSP data, as it has a better resolution (0.5 km^2), and is more sensitive, so can capture even smaller differences in night light. Each 0.5 km^2 is represented by one pixel, and the brightness (measured in nW cm^{-2} sr^{-1}) can have a value ranging from 0 to 65,535.

Data from this satellite is uploaded daily (as the VNP46A1 product) in the form of an HDF5 (Hierarchical Data Format) file. For easier data management, the entire earth is split into a 36 x 16 grid. Each tile in this grid is a square that covers a latitude and longitude of 10° (as shown in Fig 1).

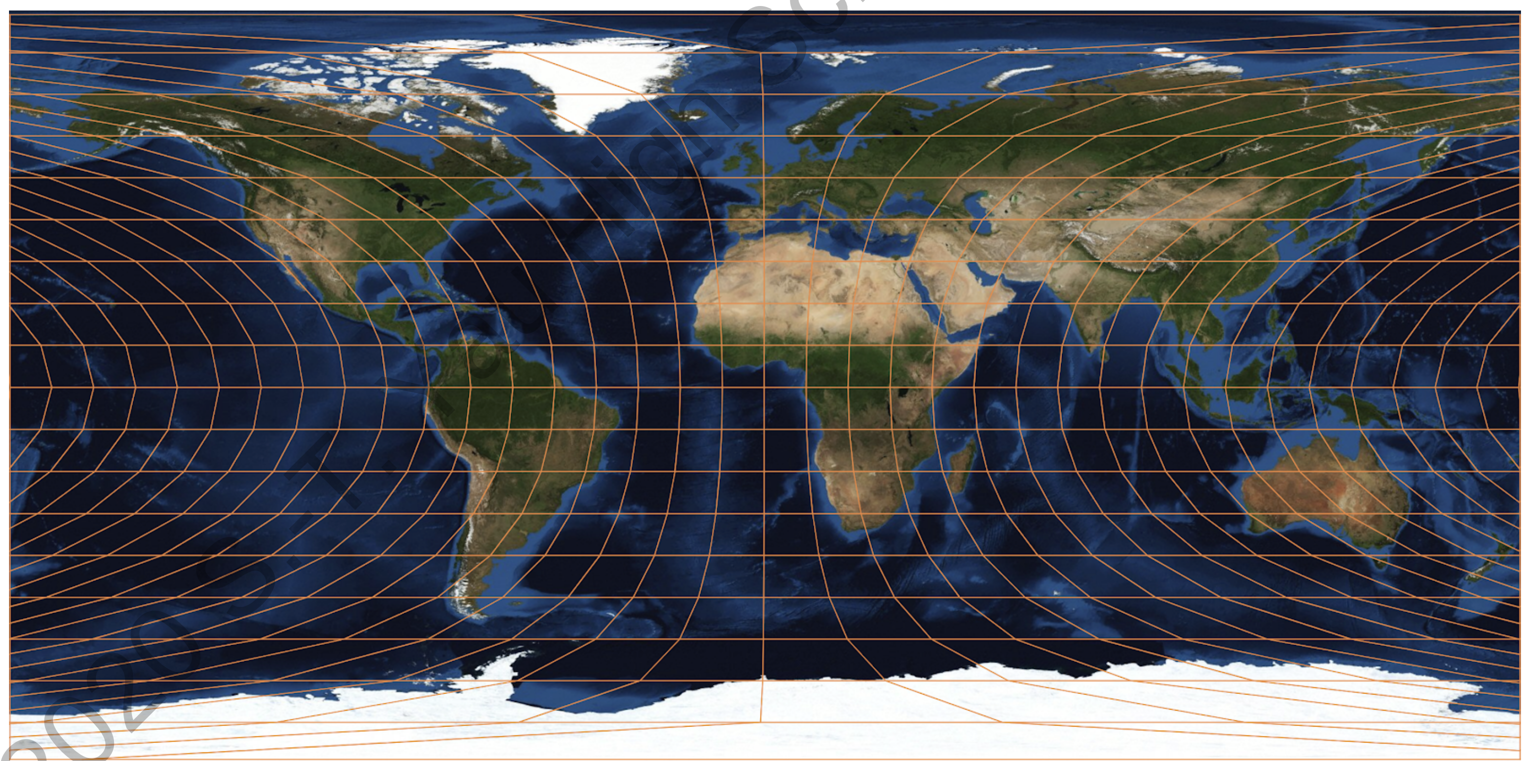


Fig 1 - Layout of Tiles on the World Map for VIIRS data

This data is also uploaded daily on massive scales, as each day's worth of data is roughly 20GB (which translates to 7 Terabytes of data annually). However, raw data collected by the sensor cannot directly be used as it contains a lot of noise from clouds and moonlight. Hence, the HDF5 file contains additional metadata, such as cloud detection, moonlight angle and reflectance, to help in correcting for these errors and arrive at a processed clean data file. This data can also be encoded with geo coordinates and converted to a GeoTIFF file which can be overlaid on a map to see night light along with pointers for cities and roads. Processing this data to help in the analysis would require an end-to-end data processing pipeline, with the ability to handle massive amounts of data, with minimal human intervention.

3.2 Electricity Consumption

Electricity consumption data was taken from India's Power System Operation Corporation Limited (POSOCO). POSOCO is a government-owned entity under the Ministry of Power which was formed in 2009 to handle all the power functions of the country. POSOCO collects data from 5 regional Load Dispatch Centres and a National Load Dispatch Centre, through which it collects data from every state daily (measured in Million Units / MU), and publishes both daily and monthly reports in the form of PDFs.

12. ENERGY COMPARISON OF NOVEMBER 2019 vs NOVEMBER 2018

REGION	STATE	Energy Requirement (MU)					Energy Met (MU)				
		Nov-18	Nov-19	Difference	% Change	Average MU/day for Nov-19	Nov-18	Nov-19	Difference	% Change	Average MU/day for Nov-19
NR	Chandigarh	93	101	8	9	3	93	101	8	9	3
	Delhi	1839	1883	44	2	63	1838	1883	45	2	63
	Haryana	3388	3367	-21	-1	112	3378	3367	-11	0	112
	Himachal Pradesh	795	822	27	3	27	793	821	28	4	27
	J&K(UT) and Ladakh(UT)	1532	1466	-66	-4	49	1242	1186	-56	-5	40
	Punjab	3394	3050	-344	-10	102	3394	3050	-344	-10	102
	Rajasthan	6738	6341	-396	-6	211	6738	6341	-396	-6	211
	Uttar Pradesh	8104	7972	-132	-2	266	8098	7934	-164	-2	264
	Uttarakhand	1015	1002	-13	-1	33	1013	1002	-11	-1	33
	NFF/Railway	82	71	-11	-13	2	82	71	-11	-13	2

Fig 2 - Monthly Electricity Consumption India Report November 2019

Fig 2 is a snapshot of POSOCO's monthly electricity consumption report for the northern region of India. We used the data from the last column in the table: the average energy met (measured in MU/day).

To also automate the process of reading data from all PDFs so that the system could be practically run for all historical data, we implemented a parsing algorithm to scan the PDFs and extract the energy met values for all states in the report.

3.3 GDP

For GDP data, we used data from the same year across cities. We found that 2018 was a very well documented year in terms of GDP data, and since it is relatively recent we chose this to be the benchmark year for all cities for regression with Night Light and Electricity Consumption data.
Moreover, we believe that an increase or decrease in night light or electricity consumption is more related to the purchasing power of an economy and hence we used GDP (Purchasing Power Parity), rather than nominal GDP, as it is a better indicator of economic activity adjusted for purchasing power across countries. Throughout the rest of the paper, GDP (PPP) will be referred to as GDP.

Furthermore, we also noticed that there was a slight discrepancy in GDP numbers, for the same city in the same year, across sources. For this reason, we used data from “Statista”, a major business data platform that we found reliable as it has been cited by researchers from many top universities and research-based companies.

Finally, for state-wise GDP data for India, we used the official government GSDP (Gross State Domestic Product) report as released by Ministry Of Statistics and Programme Implementation (MOSPI), which contained GSDP data from the year 2013 till 2018 in real terms (i.e. adjusted for inflation).

Chapter 1

4. Statistical Framework for Using Night Light to Estimate GDP Decline

4.1 Data Manipulation Pipeline

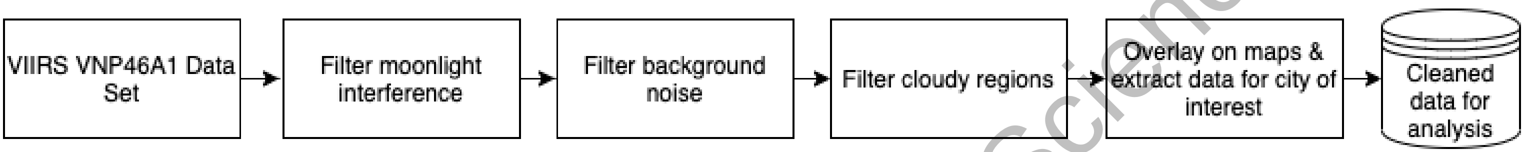


Fig 3 - Flowchart Summarizing our Data Manipulation Pipeline

4.1.1 Filter Moonlight Interference

Raw Night Light Data produced by the satellite has a lot of noise that has to be corrected and cleaned before using. One of the main issues is moon reflectance. Water bodies and land both reflect a large portion of the moonlight that falls on them, which interferes with the Night Light caused by human activity because regions which were originally dark get illuminated, showing an artificial increase in night light. Since this doesn't correlate with economic activity, we need to remove this reflectance to minimize underestimations or overestimations when measuring Night Light.

Moonlight interference is the least around New moon nights hence, we used data only from dates that were 3 days before and after a new moon night, resulting in 7 days of data. During this week, moon reflectance is minimized and almost eliminated, giving clear results of only night light. We used this approach when downloading data for both analyzing the impact of COVID and downloading data to empirically show the GDP-Night Light correlation.

4.1.2 Removing Background Noise

Also, small amounts of light are reflected by vegetation and thin clouds on isolated land. However, the light reflected by these is very low and can be easily isolated from real lighting due to the difference in brightness. We found that pixels with a brightness <30 nW cm^{-2} sr^{-1} were generally noise from such vegetation. Hence, we filtered out all pixels that had a brightness <30 nW cm^{-2} sr^{-1} to be able to better isolate lights caused by human activity.

4.1.3 Cloud Masking

Another issue with raw Night Light data is cloud diffusion. Large clouds often diffuse the light underneath them, which disperses the light that the sensor receives. Often, this causes excessive illumination in some spots, which can lead to an inaccurate result. To correct this, the VIIRS satellite data also includes information on whether a cloud was present at a given location.

Using this data, we generated a cloud mask to isolate regions where no clouds are present. However, this would lead to some gaps in the data. To fix this, we averaged data over a week, so that the gaps of one day could be filled in by data of another day. Averaging over a week also reduces noise for background noise pixels that may have a brightness more than 30, which can further improve results.

4.1.4 City to Lat Long Mapping and Data Extraction

The VIIRS night light data comes in tiles, with each tile covering an area of 10° latitude and 10° longitude. However, for analyzing the impact of COVID, we wanted to focus on segmented locations, such as a city or a road. To do this, we drew bounding boxes around the city of interest and checked the latitude and longitude of the top left and bottom right corner of the bounding box. With this, we had enough information to recreate the bounding box and extract only the pixels within this boundary.

We then extracted the information of the upper left corner of the tile and then translated the latitude and longitude information of the bounding box into pixel locations, with which we were able to extract only the area of interest. This way, given the geo-coordinates and size of any city, we can extract night light data only for that particular region.

4.1.5 Time Series Analysis

To analyze the impact of COVID, we found it best to follow with a time series analysis. By analyzing the Night Light over consecutive months, we were able to measure the impact of the virus as it progressed through different stages in a city. Moreover, by referring to the timeline of the city and checking for key dates, such as the start or end of a lockdown, we can also independently analyze how each event had an impact on the economy as well as the number of cases.

4.1.6 Source Code

Given we were processing massive amounts of data, we needed to build a fully automatic software pipeline to execute various steps of the process, such as downloading data for any range of dates, removing background noise and handling cloud interference. We modularized our code so that each part of the pipeline could be executed independently. Due to this structure, our pipeline could easily be configured to include relevant cleaning procedures and satisfy a variety of use cases. We also open-sourced our code (GitHub Link to Source Code) so that it can be easily accessed for free by other researchers.
Alternate Link to Source Code: https://github.com/Yushgoel/Nightpy

4.2 Night Light as a Measure of GDP

4.2.1 Intuition

Plotting the Night Light data for the entire world on a single night on a map gives us an intuitive understanding of the relationship between Night light and GDP.

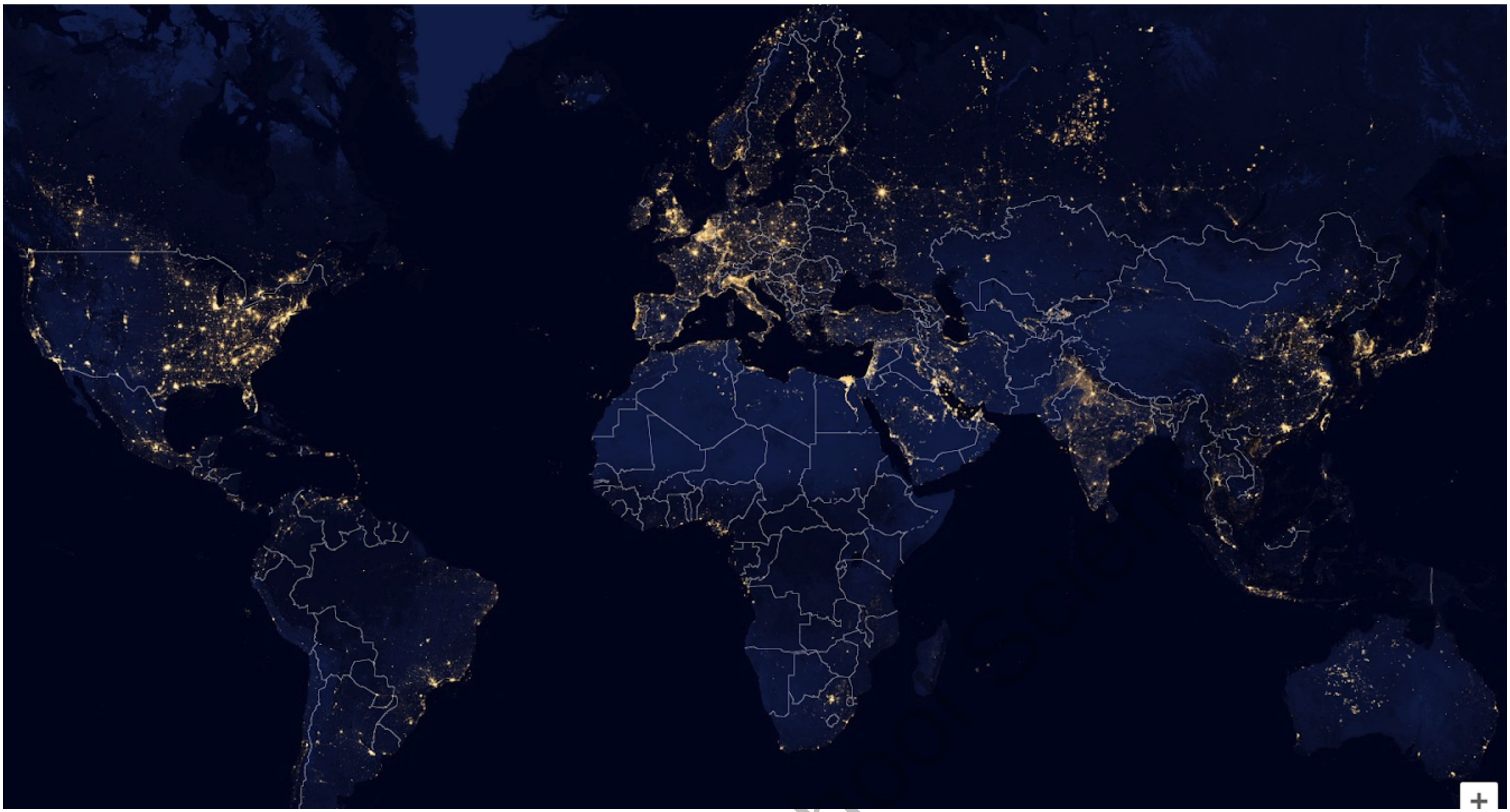


Fig 4 - Night Light For the Entire Globe[Reference 20]

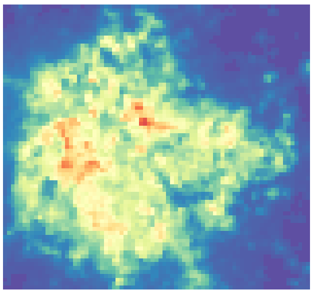
Here we can observe the following points:

1. The Eastern Portions of the U.S. are more economically developed, and we can see that they indeed have a much higher brightness than the rest of America. Also, in the Western Regions of the U.S., we know that California (lower left of the U.S.) is the most developed, and we can see that the brightness there is much higher as well.
2. We can also compare the GDP of countries visually. For example, European countries (such as Britain) have a higher GDP than many Asian countries such as India, and we can clearly see the difference in brightness here as well.
3. Within India, we can see that the Northern Regions have a higher brightness, which reflects the fact that the Northern States are more economically developed.
4. Within China, again we can see the distinction between Eastern and Western China, and we can see that major cities like Shanghai and Hong Kong are quite prominent.

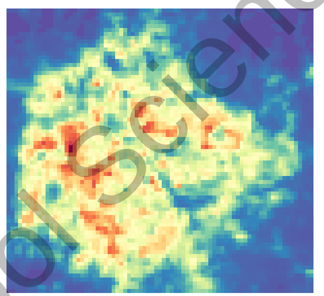
5. Within Egypt (Northern Africa), we can see that the regions near the Nile River are the most economically active, and follow the distinctive shape of the Nile River.

4.2.2 Intuition for Time Series Correlation

The correlation between Night Light and GDP is not restricted to comparing the GDP of different cities, but can also be used to see the change in GDP of a given city over time. For example, below are two images of Night Light of Bangalore - the tech hub of India.



Bangalore 2013



Bangalore 2018

In this same time frame (2013-2018), Bangalore had a roughly 80% increase in GDP, and the Night Lights for the two years also tells a similar story. Firstly, we can see that the size of Bangalore has significantly increased over the years (as depicted by the number of newly lit pixels), especially on the Eastern side of the city. We can also see that many regions in Bangalore have also drastically increased in intensity (seen by the increase in the number of red and dark red pixels), thus showing the increase in economic activity. Hence, from this comparison, we can intuitively see that GDP and Night Light are correlated even when taken from a Time Series perspective (i.e analyzing the economy of a city over time).

4.2.3 Empirical Approach

In our approach to determine the correlation between GDP and Night Light Intensity, we selected large and small cities across the globe, along with their GDP, and ran a Regression Analysis on it.

Process to Select Cities

To also make sure that we achieved a relationship that held for cities across the globe, we tried to include various cities from different continents and countries. We also wanted a set of different city sizes, so we included smaller tier-2 cities as well. Specifically, we chose cities that had well documented GDP data. Also, since we are more familiar with Indian cities, and wanted to follow with an analysis of the decline in Indian cities, we chose a higher proportion of Indian cities, roughly half. We finalized on 21 cities spread across various countries, including India, China, the United States, Canada, Australia, Mexico, Iran, Brazil, South Korea, Thailand and Taiwan.

Correlation and Regression

After selecting the 21 cities, we extracted their respective Night Lights for the year 2018, along with the GDP for the same year. Based on our review of literature, we decided to start with a simple Linear Regression as a baseline to establish the relationship. Hence, the GDP would be a function of the Night Light with the following relationship:

$$\psi = \theta \cdot \phi + \beta$$

Where

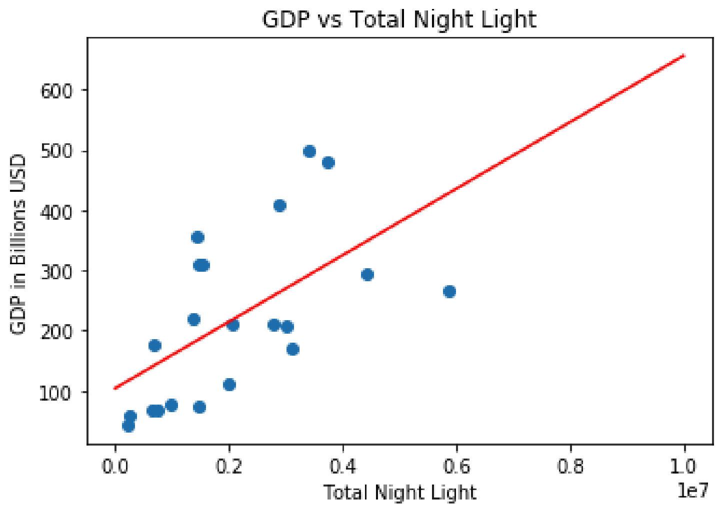
ψ is the GDP,

ϕ is the Total Night Light,

θ is the gradient of the relationship,

β is the intercept term.

Once we ran the Regression to evaluate the optimum θ and β, we plotted the equation (in Fig 5) captured by this relationship along with the data points to see how well it fit the data.

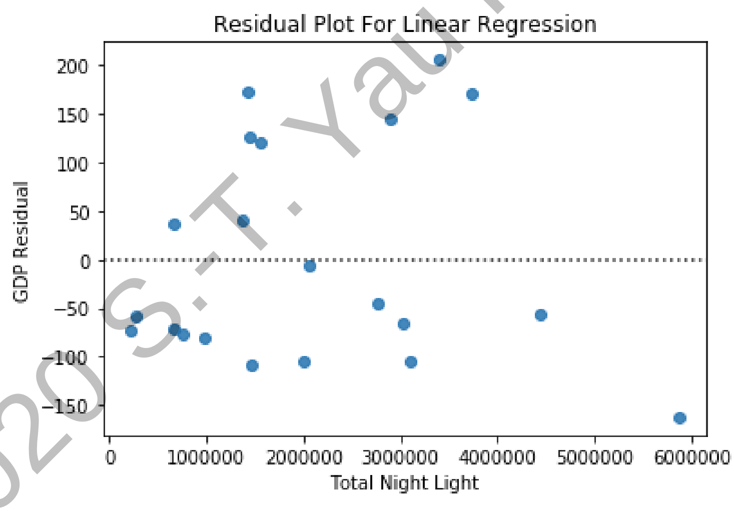


θ = 5.52e-05

β = 104.06

Fig 5 - GDP vs Total Night Light

To quantify the error in the Linear Regression's output, we also evaluated it on the Mean Absolute Error (MAE) Metric. We chose MAE instead of Root Mean Squared Error (RMSE), because of the way the two metrics penalize the fit. In RMSE, a 5 % error prediction for a city with a GDP of 500 billion USD is penalized much more than a 20% error in a city with a GDP of 100 billion USD. Since this large error for smaller cities is not desirable and we want to penalize estimation errors in large and small cities equally, we find the MAE metric more appropriate. The MAE for the standard Linear Regression Plot was 96.5. We also plotted a residual plot (Fig 6) to see the deviation of the predictions from the actual data.



However, we can also visually see that the Linear Regression equation is not able to model many points very well. This may be because the relationship isn't linear. To determine if a non-linear relationship fits the data better, we tried to run a Regression on a Log - Log plot.

Fig 6 - Residual Plot For Linear Regression

Log Log Relationship

A Log-Log regression uses logarithmic scales on both axes, i.e both the x and y axes are logarithmic. Log-log regression models an equation with the following relationship:

$$ln(\psi) = \theta \cdot ln(\phi) + \beta$$

Manipulating the equation to keep ψ as the subject, we have:

$$\psi = \phi^{\theta} \cdot e^{\beta}$$

Where

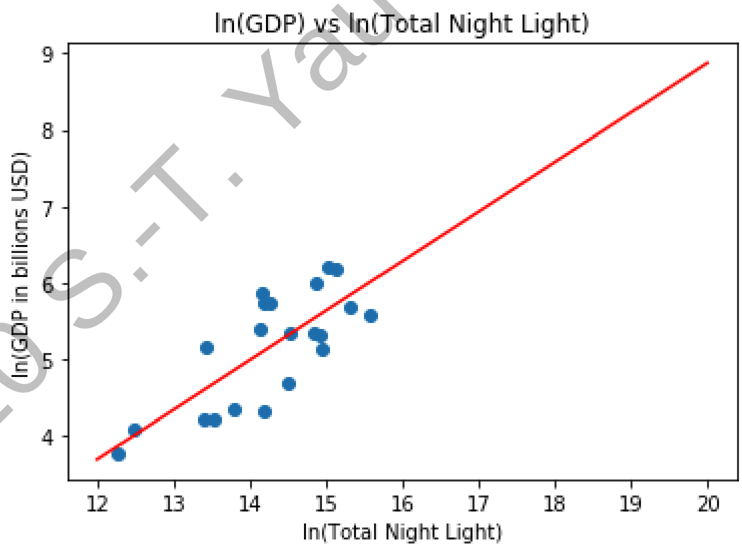
ψ is the GDP,

ϕ is the Total Night Light,

θ is the gradient of the relationship,

β is the intercept term.

Therefore, a log-log plot presents monomials (single term polynomials) as a straight line. This way, we can capture any non-linear monomial relationship accurately. Moreover, in our case, we are dealing with variables that grow large very quickly, and thus cities with even a slightly higher GDP will have a huge impact on the Linear Regression equation. A log-log plot can compress and reduce this gap, resulting in a better fit to the data. We ran the regression on the log-log plot and received the values of θ and β, we plotted the data points along with the equation.



$\theta = 0.65$

$\beta = -4.05$

Fig 7 - Log Log plot for GDP vs Total Night Light

From Fig 7, we can see that the new equation from the log-log plot is a much better fit to the data. The θ of 0.65 shows us that the relationship between the variables is not linear, rather more towards the square root. This tells us that although for small night light values, Night Light is more linearly correlated with GDP, as night light increases, the GDP increase saturates. This also makes intuitive sense. As a small city develops, it receives more lighting for the roads and new buildings which reflects more economic activity. However, beyond a level of development, most of the additional lighting is purely for recreation, and generally not for productive use. To numerically compare the log-log plot with the standard axes, we cannot directly compare the MAE for the two equations, since the deviations in the log-log plot are reduced. Hence, the predictions of the equation have to be exponentiated to bring them on the same scale as the baseline Linear Regression. The MAE for this equation was 88.33.

When observing our regression results for the log-log plot, we found that our equation was systematically underpredicting GDP for cities in developed countries, while slightly over predicting GDP for cities in developing countries. We realized that this could be due to the fact that developed economies are primarily based on the tertiary sector, and produce more economic output per unit of lighting, as compared to developing economies.

Based on this observation, we decided to add another parameter of a Boolean Flag to indicate whether a city belongs to a developed economy or not. We used a GDP per capita threshold of 45000 USD, so cities with a higher GDP per capita were marked as developed. This way, our regression can better differentiate between the behaviour of developing and developed cities and make more accurate GDP predictions. Therefore, we have our updated equation as follows:

$$ln(\psi) = \theta \cdot ln(\phi) + \gamma \cdot \omega + \beta$$

Where

ψ is the GDP,

ϕ is the Total Night Light,

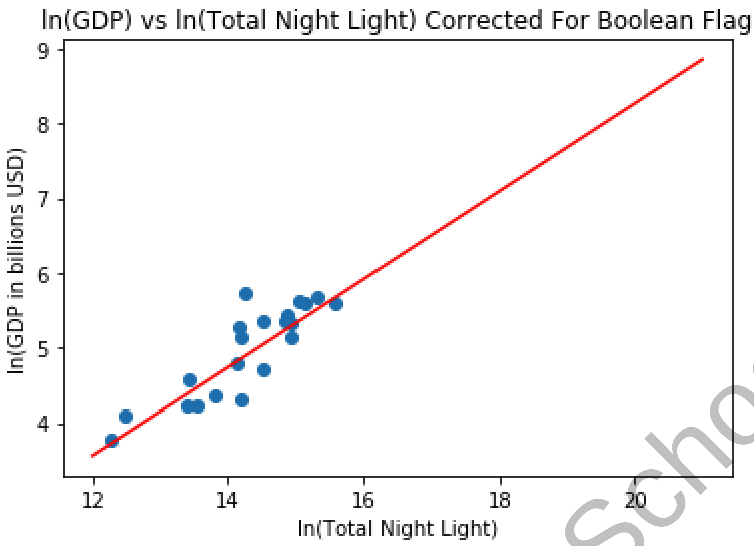
θ is the gradient of the relationship,

ω is the boolean flag indicating a developed or developing economy,

γ is the average difference in GDP between a developed and developing economy,

β is the intercept term.

We again plotted the regression line (Fig 8) but also corrected for the bool flag to accurately visualize the fit in 2 dimensions. The results of the regression gave us the following values.



$\theta = 0.587$

$\gamma = 0.807$

$\beta = -3.48$

Fig 8 - Log Log Plot for GDP and Total Night Light With Boolean Flag

After exponentiating the predictions from this new equation, we arrived at an MAE of 35.93, which is much better than previous results. This reinforces our notion that developed and developing countries use Lights differently, making it hard for a simple model (without any knowledge of which is a developed city) to fit the data. We summarized the results of all 3 models in the table below.

Regression Technique	**MAE (Mean Absolute Error)**	θ	γ	β
Linear Regression	96.5	5.52e-05	-	104.06
Log-Log Plot	88.33	0.65	-	-4.05
Log-Log Plot + Boolean Flag*	35.93	0.587	0.807	-3.48

* To distinguish between developed and developing cities

Table 1 - Summary of Different Regression Models

We also summarized the correlation between the variables. To also make sure that such correlation results were not purely accidental, we tabulated the p-value for each coefficient. P-value is a good measure for the statistical significance of results.

	GDP	ln(GDP)	p-value	Significant at p-value < 0.01
Total Night Light	0.587	-	0.005	True
ln(Total Night Light)	-	0.7623	0.000059	True
Boolean Flag*	0.679		0.000713	True

* To distinguish between developed and developing cities

Table 2 - Statistical Significance of Correlation between Different Variables

We think the log log relationship along with the boolean flag best represents the true relationship, due to its tight fit and low p-value.

4.2.4 Model to Find GDP Decline

The regression equation for our selected model (reproduced below) can then be modified to find the percentage difference in GDP over two time periods using the Night Light difference.

$$ln(\psi) = \theta \cdot ln(\phi) + \gamma \cdot \omega + \beta$$

Let

ϕ_1 *and* ϕ_2 be the Total Night Light for a place on two different time periods t_1 and t_2 respectively.

ψ_1 *and* ψ_2 be the GDP for a place on t_1 and t_2

Then, we have:

$$ln(\psi_1) = \theta \cdot ln(\phi_1) + \gamma \cdot \omega + \beta$$

$$ln(\psi_2) = \theta \cdot ln(\phi_2) + \gamma \cdot \omega + \beta$$

If we take the difference between the two:

$$ln(\psi_2) - ln(\psi_1) = \theta \cdot ln(\phi_2) - \theta \cdot ln(\phi_1)$$

Since β is a constant and $\gamma \cdot \omega$ stays fixed for a given place, the two terms cancel out in the subtraction. Manipulating it further:

$$ln(\frac{\psi_2}{\psi_1}) = \theta \cdot ln(\frac{\phi_2}{\phi_1}) \Rightarrow \frac{\psi_2}{\psi_1} = e^{\theta \cdot ln(\frac{\phi_2}{\phi_1})} \Rightarrow \frac{\psi_2 - \psi_1}{\psi_1} = e^{\theta \cdot ln(\frac{\phi_2}{\phi_1})} - 1$$

This gives us the ratio between the GDP between t_2 and t_1. Subtracting 1 from this ratio and converting to percentage, we have the percentage change in GDP between the two time periods. Using this equation, we can find the percentage decline or increase in GDP for various cities due to the COVID-19 pandemic.

4.3 Framework to Estimate GDP Decline of a city for COVID

Analysis of a few initial cities (through our data preprocessing pipeline) indicated that Total Night Light, in and of itself, didn't give an accurate picture of the decline in a city. To investigate this further, we plotted the percentage change of the city, pixel-wise, to understand why the data was not representative of the decline and saw that some parts of the city went up in brightness, while others went down. Upon further investigation, we realized that the regions that went up in lightning were mainly residential, whereas the regions that experienced a decline were mostly commercial. This observation makes sense because either due to official lockdown or self-imposed restrictions, people would tend to stay home, causing residential lighting to go up. In Figure 9, we can see the increase or decrease in pixel values for Bangalore from January 2020 to March 2020. We can see that towards the eastern outskirts of the city, there has been an increase in lighting (shown in green), which is the increase in brightness of residential areas.

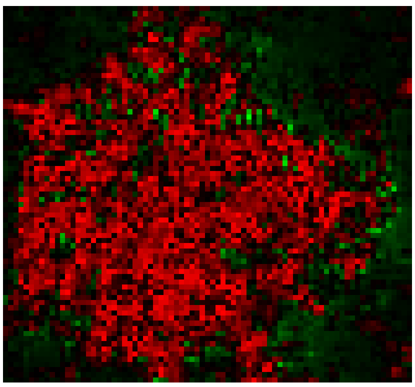


Fig 9 - Change in Brightness for Bangalore from January to March 2020

From this, we concluded that change in total night light should not be used to understand the economic impact of COVID. Hence, we needed to develop a system to distinguish between mostly residential and mostly commercial areas based on the night light radiance and only use night light from commercial areas to estimate the economic impact of COVID.

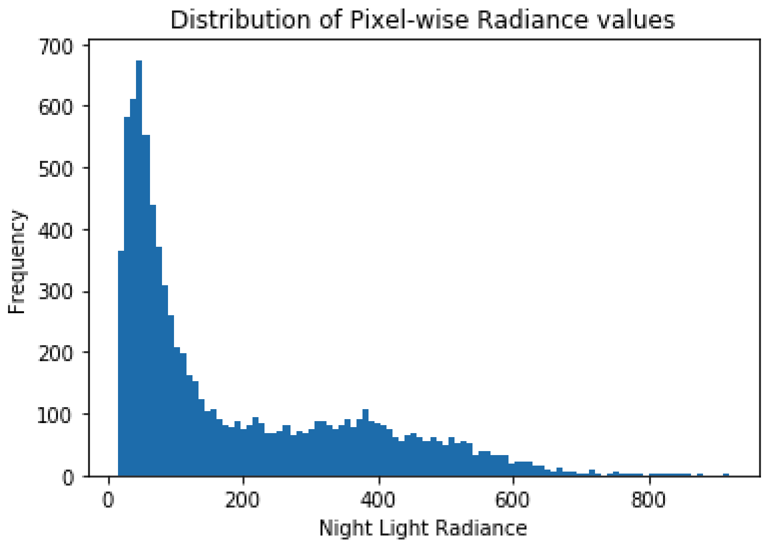


Fig 10 - Distribution of Pixel-wise Radiance Values

We think that the intensity of illumination for a given pixel should represent whether it is a mostly commercial area or a mostly residential region.

To understand what threshold of night light intensity to use to segment the data, we plotted a histogram to understand the distribution of pixel values (shown in Fig 10).

From the histogram, we can see that the city is divided into 2 parts, regions which have a relatively low brightness (but are very large in number, which is denoted by the hump at low values of Night Light Radiance), and regions with a much higher brightness (but lower in number, forming a long tail towards the end of the histogram). To differentiate between commercial and residential areas, we need to distinguish between the hump and the long tail, which appears at a radiance value of 190 nW cm^{-2} sr^{-1}.

Hence, we took values larger than 190 nW cm^{-2} sr^{-1} as mostly commercial, and regions with a lower brightness as mostly residential areas. One limitation for this approach, however, is that some regions may have a mix of residential and commercial areas, especially in developing cities, where a lot of economic activity is mixed rather than segmented. Due to this, an increase in the residential portions may undermine the decline in the commercial regions, leading to a slightly inaccurate result. However, due to the relatively good resolution of a city (1 pixel = 0.5 km^2), such regions will be small in number and shouldn't cause too much of a problem.

However, isolating commercial regions month-wise, using this threshold, is not appropriate. The reason is that as commercial regions are shut down and lose brightness, they may go to an illumination range less than 190 nW cm^{-2} sr^{-1} and won't be counted as commercial regions anymore. Hence, some areas will be "lost" in our calculations, causing an overestimate of the decline. To fix this, we implemented a benchmarking month, i.e. a month in which we can safely assume that all economic activity was normal. We then identify commercial regions from this month using the threshold of 190 nW cm^{-2} sr^{-1}. For all subsequent months, we isolate all the pixels which were flagged as commercial from our benchmarking month and calculate the decline only from these pixels. This way, we avoid pixels being "lost" and receive a more accurate estimate.

5. Application of Framework to Impact of COVID-19

5.1 Estimated Economic Impact of COVID-19 on cities

Proceeding with our fully developed pipeline and framework for understanding COVID on a city, we quantitatively analyzed the economic impact of COVID-19 on 12 cities of interest.

% Change in GDP From Benchmark Month

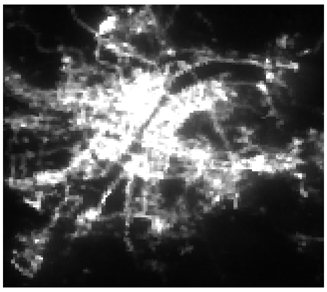
	India									United States		China
	Bangalore	**Salem**	**Mumbai**	**Pune**	**Delhi**	**Ahmedabad**	**Surat**	**Jaipur**	**Bhopal**	**New York**	**Miami**	**Wuhan**
Dec	-	-	-	-	-	-	-	-	-	-	-	0.0
Jan	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-46.6
Feb	-1.2	0.7	-1.6	-0.1	-5.4	5.6	0.3	1.4	3.4	-6.9	-8.9	-35.8
Mar	-10.0	-22.1	-11.7	-14.2	-10.4	-23.6	-11.2	-29.4	-24.0	-12.0	5.6	-45.2
Apr	-17.3	-37.7	-8.6	-2.3	-11.0	-7.1	-8.4	-23.2	-19.1	-19.3	-21.3	4.6
May	-24.1	-34.3	-13.7	-8.9	-9.7	-5.8	-9.6	-7.6	-12.1	-17.3	-24.8	-4.0

Table 3 - Summary of GDP decline Obtained Using Framework

As we can see from table 3, different cities had different levels of economic decline, having a wide range of peak impact, from 1% to 46% decline. There would have been several factors contributing to the variance in decline, and we explore some of these factors in this section.

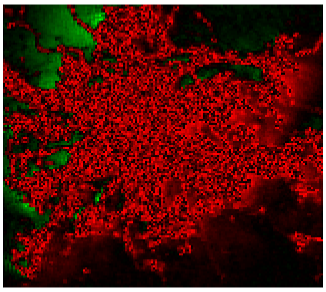
Wuhan, China

The first case in Wuhan was officially declared in mid-December. Initially, this was thought to be a case of pneumonia and was only confirmed as the coronavirus in January, and by then the government started taking action. Since the economy was operating at a normal level in December, we took it as the benchmark month. By mid-January, the virus had spread rapidly across the region, and the government imposed an extremely strict lockdown, essentially shutting down the entire economy. This is also reflected in the night light values, through which we estimate a 46.6% decline in GDP. This can also be visually seen in the figure below.

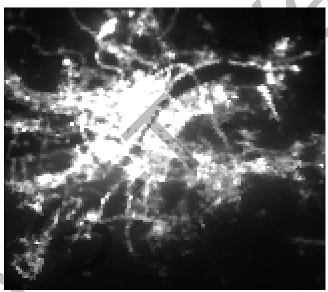


Wuhan December

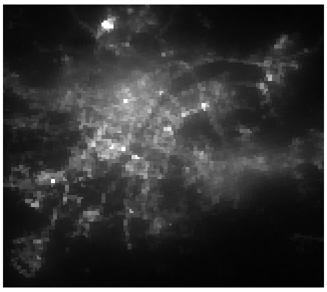
Wuhan January



Diff Jan - Dec



Wuhan April

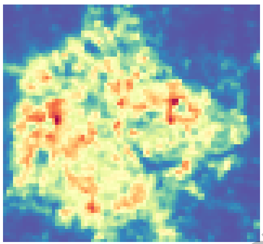


Wuhan June

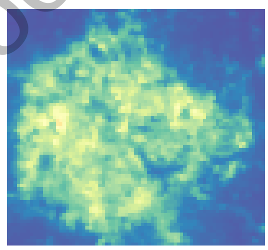
As seen from the difference plot above, Wuhan suffered an incredible decline across the city, indicating both the seriousness of the lockdown and the resulting large economic decline. Reports show that the situation was completely controlled, and the lockdown was opened by April, and night light data also shows that Wuhan's economy went completely back to what it was in December. However, the data also shows that in June, Wuhan's economy went down again to about 62% of what it was in December. This may be a self-imposed or loose lockdown, in light of a potential second wave.

Bangalore, India

In January, India hadn't reported even a single case of the coronavirus, so we can safely assume that there was no decline in economic activity in this month. Hence, we chose this as the benchmark month for Bangalore. We know that the government imposed a nationwide lockdown in India on 25th March, and the Night Light shows a 10% decline in GDP in that month. However, in the following months, as the lockdown was more strictly enforced, the GDP dipped further, having a 17% decline in April, and a 24% decline in May.



Bangalore January



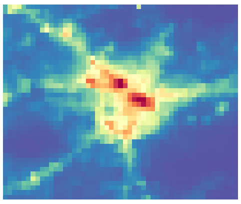
Bangalore May

Another reason for the continuous decline might be due to the counter-urbanization of migrant workers. In around April - May, it was reported that migrant workers from nearby towns and villages started leaving the city, due to lack of jobs. This may have been the reason behind the further decline in May with respect to April.

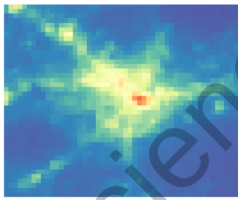
Salem, India

Salem is a tier 2 city in India, which is heavily reliant on industries (textile and steel) as well as highway traffic (since it is a hub for three national highways) for its economy. The city went into lockdown with the rest of the nation, so it had the exact same timeline as Bangalore. By March (when the lockdown started), Salem experienced a GDP decline of 22%. In April (when the lockdown was enforced for the entire month), the decline increased to 38%.

Salem might have seen such a drastic decline since Steel and Textiles are generally used as raw materials in other industries. As many of these consumer industries reduced their production, the demand for steel and textiles dropped sharply. Hence, steel and textile industries in Salem might have been shut down for a longer duration. By May, the economic decline stabilized at 34%, indicating that there was no further tightening of policies.



Salem January



Salem May

Section 5.2 (impact on intercity transport) shows that there was a 38% decline in traffic between January and May 2020. Since Salem is also a hub of three national highways with large economic dependence on interstate transportation, it is not surprising that it had a similar GDP decline of 37%. Hence, using night light, we were even able to capture GDP changes for smaller Tier 2 cities, without compromising on precision.

New York and Miami. U.S.

In early January, The U.S hadn't reported any cases, so again we assumed that there was no decline in economic activity in this month and chose this as the benchmark month. New York and Miami didn't have any enforced lockdowns until April and had only a partially imposed lockdown during this month. Hence, most of the economic decline would have been due to self-imposed isolation by people. In February, New York and Miami had a GDP decline for roughly 7.9%. This early decline in February (compared to Bangalore) is reasonable since the U.S had a much higher number of cases during this time than India. By April, New York and Miami had a 20.3% decline in GDP. This is the same month during which the two cities had a partial lockdown enforced. However, by June, once the cities reopened, the economic activity slightly revived by 5.3% from April.

We expected to see a smaller decline in US cities because the lockdowns enforced were partial and for short periods of time, as compared to the strict long-duration lockdowns imposed in India and Wuhan.

In our estimates, the average decline for New York and Miami (using Night Light Data) for Q2 2020 (Apr - June) was roughly 12.3%. Incidentally, on July 30th 2020, The U.S. Bureau of Economic Analysis reported a 7.2% decline for the quarter. Given that New York was the epicentre of early COVID growth in the U.S., and that Miami's economy is heavily based on leisure and tourism, it is expected that these two cities would have had a higher decline as compared to the average of the entire country. This gives further credence to our results and demonstrates that Night Light can give reasonable estimates for GDP decline in real-time.

5.2 Impact of COVID-19 on Inter-City Transport

Night Light on roads is primarily emitted by traffic that passes by, and any decline in this Night Light would be a direct result of the decline in traffic. Hence, for understanding the impact of COVID on such intercity transport, we would pick major highways and measure the decline in luminosity.

Our intercity research was mostly based on Indian cities. We picked major metropolitan cities of India with strong trade connections and then picked highways connecting them. Here, we focused on highways officially registered as a National Highway (NH) or an Asian Highway (AH) so that we were confident that the road was purely an intercity road rather than a branch passing through a small town.

However, we also noticed that when following NH or AH roads, there were many villages or towns that surrounded the road. Moreover, many of these towns were less than 500 m in size, so with the given satellite data resolution, we couldn't distinguish such villages from the road. Also, the width of the road was fairly narrow, compared to the 500 m resolution of the data. Therefore, to ensure that any decline in night light measured on the roads was due to traffic decline and not economic decline in a nearby village or town, we isolated segments of road that were surrounded by forests or uninhabited land using Google Satellite imagery. In these segments, we knew that all of the light was emitted from traffic (ignoring the small constant amount emitted by street lights), so the change in measured night light should be directly proportional to the change in intercity traffic along these segments of road.

	Bangalore to Kanyakumari (AH 45)	**Bombay to Delhi Road (NH 8)**	**Bangalore to Hyderabad (NH 7)**
Jan*	0.0%	0.0%	0.0%
Feb	-10.94%	6.16%	-6.88%
Mar	-23.52%	-36%	-28.34%
Apr	-41.48%	-28.7%	-43.72%
May	-28.15%	-39.33%	-50.61%

* Used as benchmark month

Table 4 - Decline in Traffic for Intercity Highways

We can see that by March, there was roughly a 30% decline in intercity logistics. However, since the lockdown started on 25th March, we only saw the partial effect of the lockdown on our traffic decline. The full effect of the lockdown can be seen in April, where roads had an average decline of 37% in night light radiance caused by a decline in traffic. This decline continued throughout May, which saw an average 39% decline in light due to traffic, and hence also 39% decline in intercity traffic.

5.3 Government Response to COVID and Resulting Economic Impact

To quantitatively compare the impact of different government policies on the economy, we first segregated the responses of different countries into broadly three categories: those that imposed a full-scale strict lockdown (mainly India and China), those that imposed a partial lockdown (shutting down few services, closing schools and some institutions) and no lockdown (minimal government intervention, mostly self-imposed isolation). Using this system, we divided all of the cities we analyzed into these three buckets. We then took the economic decline measured by night light 4 months after the benchmark month (which was adjusted from city to city accordingly) and then averaged this number to get an estimate of average economic decline for each category. Table 5 summarizes our results:

Strict Lockdown		**Partial Lockdown**		**No Lockdown**	
City	**%Decline**	**City**	**%Decline**	**City**	**%Decline**
Mumbai	14%	Ahmedabad	7%	Seoul	3%
Bangalore	24%	Surat	11%	Taipei	0%
Pune	14%	Bhopal	12%	Tokyo	11%
Wuhan	47%	New York	18%		
Mysore	53%	Miami	11%		
Average	30.4%	Average	11.87%	Average	4.81%

Table 5 - Impact of Government Policies on GDP Decline

Thus, we can see that there is a significant difference in the economic decline based on the type of lockdown imposed. Places with strict lockdown experienced an average of 18.5% more GDP decline than places with partial lockdowns and 25.6% more than places with no lockdown. Also, places with a partial lockdown experienced an average of 7.06% more decline than places with no lockdown. The economic decline in places with no lockdown can be explained by self-imposed isolation in these regions. These results confirm our intuition, as places with strict lockdowns had a large economic decline, and places with a partial lockdown didn't decline as much, but declined more than places with no lockdown.

These results quantify the incremental decline in economic activity for different strictness levels of the lockdown. This can be used by policymakers to make informed tradeoffs between economic decline and the cost borne by society due to stress on the healthcare system when making decisions on lockdowns in the future.

Chapter 2

6. Statistical Framework for Using Electricity Consumption To Estimate GDP Decline

6.1 Electricity Consumption as a Measure of GDP

6.1.1 Intuition

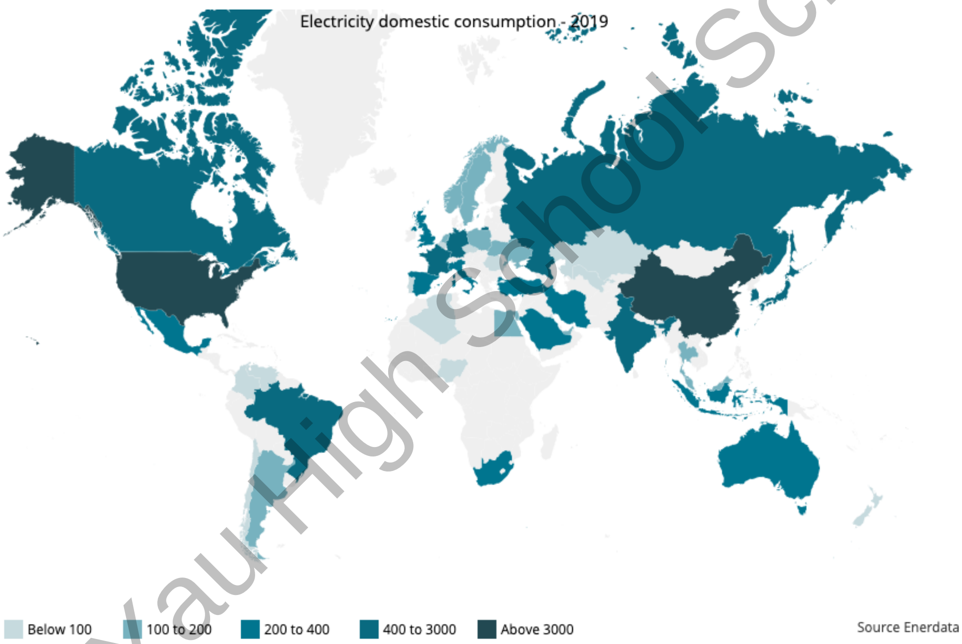


Fig 11 - Country Wise Electricity Consumption

The map above is a pictorial representation of the amount of electricity that was consumed by all the countries around the world in 2019 (in TWh), which gives us an intuitive sense of how electricity consumption correlates to GDP. In the map, the countries with the highest electricity consumption are the USA and China, which are also the countries with the highest GDP in the world. Moreover, when looking at the countries in the next range of electricity consumption, we see India, Russia, Japan, South Korea, Canada, Brazil, Germany, France, Mexico, United Kingdom etc, which are all countries in the list of the top 20 highest GDP's in the world. Looking at countries in the below 100 TWh category, the same trends can be seen. Countries such as Nigeria, Uzbekistan, Colombia, Romania etc. fall in this category, and these also happen to be the countries at the lower end of the GDP spectrum.

6.1.2 Empirical Approach

Our expectation is that different countries would have significantly different behaviour in terms of electricity consumption due to various factors such as the culture of the country, as well as it's location, since colder regions would have a higher electricity consumption used for heating. Hence, we decided to focus on one homogenous dataset and picked a single country which would largely remove the cultural effects and weather variations. Due to familiarity of the authors and the accessibility of a reliable data source, we decided to primarily focus on India. We used the 2018 GSDP data (measured in ₹ 10 millions) and the average Million Units per day for all 29 Indian States, and used this for the regression analysis.

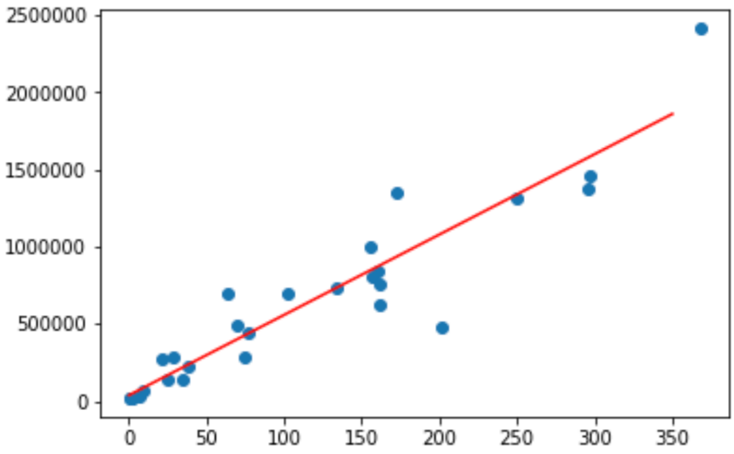
Linear

Our initial approach to model the data was linear regression, as a baseline to establish the relationship. Hence, similar to Night Light, the GDP would be a function of the Electricity Consumption with the following relationship:

$$\psi = \gamma \phi + \beta$$

Where
ψ is the GDP,
ϕ is the electricity consumption,
γ is the gradient of the relationship,
β is the intercept term.

Once we ran the regression to evaluate the optimum values of γ *and* β, we plotted the equation captured by this relationship (Fig 12) along with the data points to see how well it fit the data.



γ = 5202.82
β = 37201.8

R(Correlation Coefficient) = 0.94

Fig 12 - Electricity Consumption vs GDP

We did not think that the data modelled well enough since there are many outliers, and only a few data points seem to be on the line of best fit. To quantify how well this relationship fits the data, we chose the MAE metric, for the same reasons as Night Light). The MAE for this model was 128940.65, which is 21% of the mean. We also quantified the correlation between the two, and arrived at 0.94.

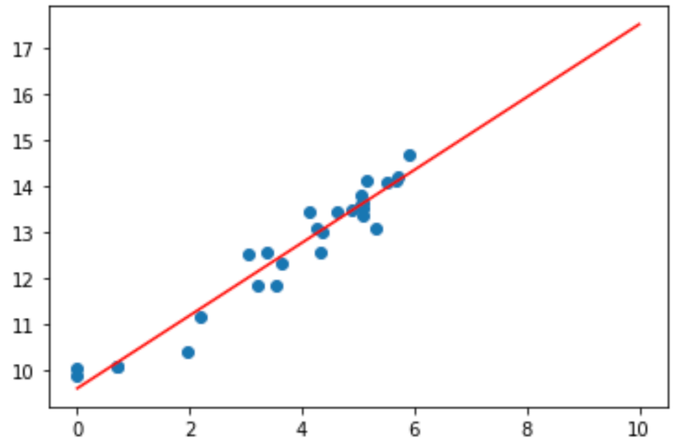
While observing the graph, we also observed that potentially, a non-linear regression would fit the data better, and increase the correlation even further. Using a log-log plot would give us this polynomial relationship with the exact degree, using which we could establish a better relationship..

Log-log scale

As explained before, a log-log plot uses logarithmic scales on both axes. The equation that models this relationship is as follows:

$\ln(\psi) = \gamma * \ln(\phi) + \beta$

Where
ψ is the GDP
ϕ is the electricity consumption,
γ is the gradient of the relationship,
β is the intercept term



γ = 0.79
β = 9.61
R = 0.97

Fig 13 - Log Log Plot of Energy Consumption Vs GDP

Compared to the previous model, there are a smaller set of outliers and all the observations seem much closer to the line of best fit. This is also confirmed with the Pearson correlation coefficient value of 0.97 which suggests an extremely strong correlation between electricity consumption and GDP. The MAE for this relationship was 132916.67, which is roughly 22% of the mean of all observed values, showing that we didn't sacrifice too much on this metric, but improved the correlation.

To summarise:

Regression Technique	**MAE**	γ	β
Linear Regression	128940.65 (21% of mean)	5202.82	37201.8
Log-Log Plot	132916.67 (22% of mean)	0.79	9.61

Table 6 - Summary of Regression Techniques

To make sure that this relationship was not due to spurious regression, we checked the statistical significance of our results using the p-test.

	p-value	**Significant at p-value < 0.01**
Electricity consumption vs. GDP	<.00001	Yes
ln (Electricity consumption) vs. ln (GDP)	<.00001	Yes

Table 7 - Statistical Significance of Correlations

6.2. Framework to Estimate GDP Decline of a State

The regression equation for our selected model (reproduced below) can be modified to find the percentage difference in GDP over two time periods using the change in electricity consumption during this time period.

$$ln(\psi) = 0.79\, ln(\phi) + 9.61$$

Let ϕ_1 *and* ϕ_2 be the total electricity consumption for a certain state in time periods t_1 and t_2 respectively. And ψ_1 *and* ψ_2 be the GDP for a place on t_1 and t_2.

Then, we have:

$\ln(\psi_1) = \gamma * \ln(\phi_1) + \beta$

$\ln(\psi_2) = \gamma * \ln(\phi_2) + \beta$

If we take the difference between the two:

$\ln(\psi_2) - \ln(\psi_1) = \gamma * \ln(\phi_2) + \beta - (\gamma * \ln(\phi_1) + \beta)$

Considering that β is a constant, it would get cancelled in the subtraction. Manipulating it further:

$$ln\left(\frac{\psi_2}{\psi_1}\right) = \gamma \cdot ln\left(\frac{\phi_2}{\phi_1}\right)$$

$$\therefore\ ln\left(\frac{\psi_2}{\psi_1}\right) = ln\left(\frac{\phi_2}{\phi_1}\right)^{\gamma} \Rightarrow \frac{\psi_2}{\psi_1} = \left(\frac{\phi_2}{\phi_1}\right)^{\gamma} \Rightarrow \frac{\psi_2 - \psi_1}{\psi_1} = \left(\frac{\phi_2}{\phi_1}\right)^{\gamma} - 1$$

By subtracting 1 and multiplying by 100, we can calculate the percentage change in GDP from t_1 to t_2.

6.2.1 Difference in Differences Approach

When using the above framework to estimate the GDP change, we are generally interested in the GDP change that can be directly attributed to some major event, such as a lockdown in response to the COVID-19 pandemic, so a difference in differences framework would be more appropriate in this scenario. More specifically, we assume that the previous years represent the natural patterns in electricity consumption across the time periods t1 and t2 (the same time period but a few years ago; for example, if t1 represents January 2020, and t2 represents March 2020, then we would take January 2019 and March 2019 as a baseline for natural GDP change. Note that the year taken as the baseline is flexible and should be chosen when there is no major economic event going on). Then, any change in the time period t_1 and t_2, which is not seen in the baseline year, would be attributed to the economic event of interest (in our example, any extra decline seen would be attributed to lockdowns).

We have:

$$\frac{\psi_2^\tau - \psi_1^\tau}{\psi_1^\tau} = \left(\frac{\phi_2^\tau}{\phi_1^\tau}\right)^{\gamma} - 1$$

Where $_1^\tau$ and $_2^\tau$ refer to the same time period t_1 and t_2 τ years ago.

Then, the GDP change attributed (extra GDP change) to the event is as follows:

$$\frac{\psi_2 - \psi_1}{\psi_1} - \frac{\psi_2^\tau - \psi_1^\tau}{\psi_1^\tau} = \left[\left(\frac{\phi_2}{\phi_1}\right)^{\gamma} - 1\right] - \left[\left(\frac{\phi_2^\tau}{\phi_1^\tau}\right)^{\gamma} - 1\right]$$

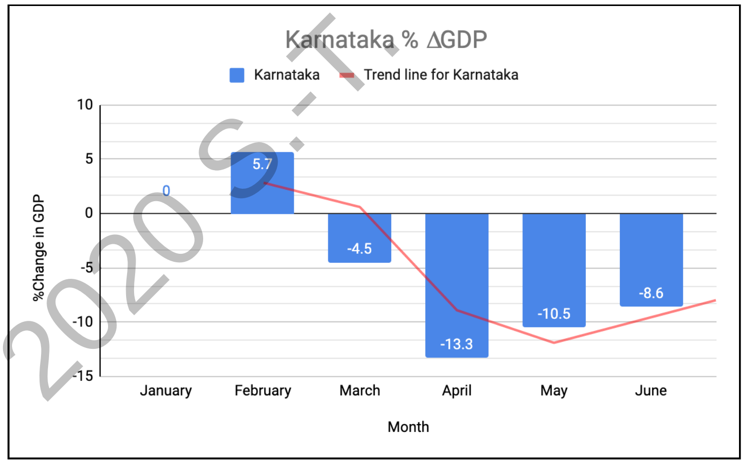
This equation will take into account the difference in differences methodology, and will give us the GDP decline attributed to the event of interest. In our scenario, we will be focusing on the recent COVID-19 pandemic. Since there were no major events in 2019 that could affect GDP, we chose $\tau = 1$ and took t_1 as January 2020 throughout the analysis.

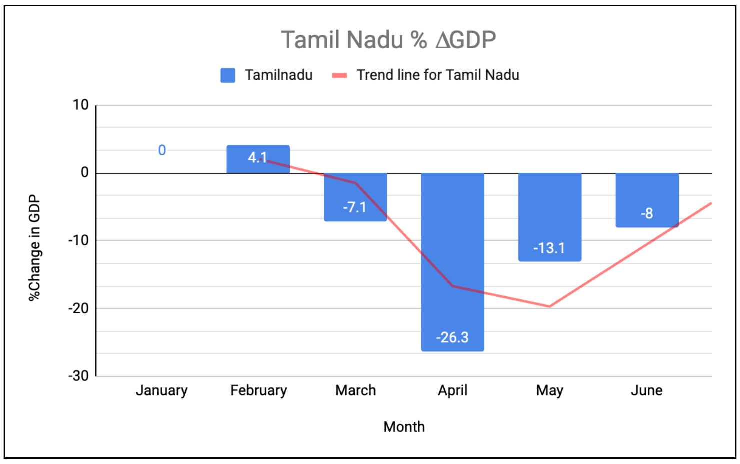
7. Application of Framework to Impact of Covid-19

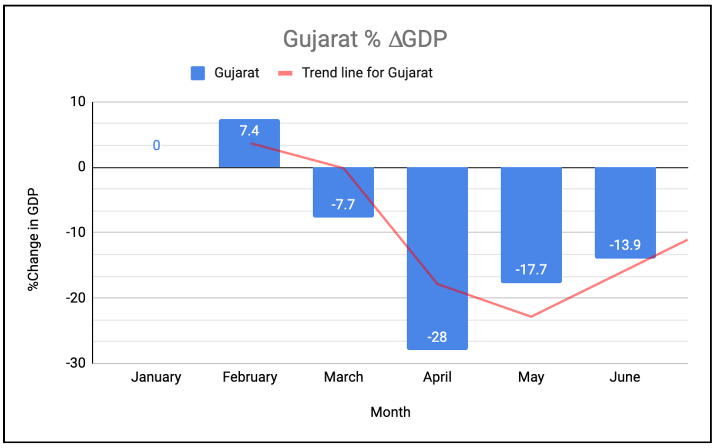
Using the framework detailed above, we calculated the estimated decline in GDP for every state from January to June of 2020.

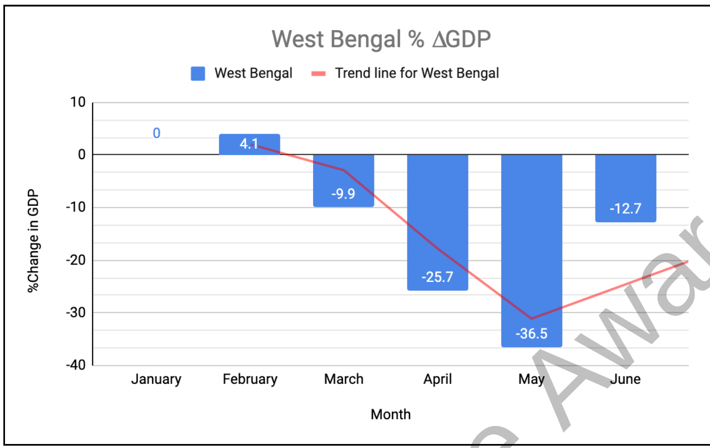
	January	February	March	April	May	June
Karnataka	0.0	5.7	-4.5	-13.3	-10.5	-8.6
Tamilnadu	0.0	4.1	-7.1	-26.3	-13.1	-8.0
Maharashtra	0.0	6.4	-8.4	-24.4	-19.1	-22.5
Andhra Pradesh	0.0	5.1	-7.8	-17.2	-10.5	-11.3
Gujarat	0.0	7.4	-7.7	-28.0	-17.7	-13.9
West Bengal	0.0	4.1	-9.9	-25.7	-36.5	-12.7
Madhya Pradesh	0.0	14.1	-4.2	-11.7	-5.0	-9.2
Rajasthan	0.0	6.5	-11.6	-16.0	-4.9	-4.7
Uttar Pradesh	0.0	6.2	-14.9	-21.2	-23.2	-10.5

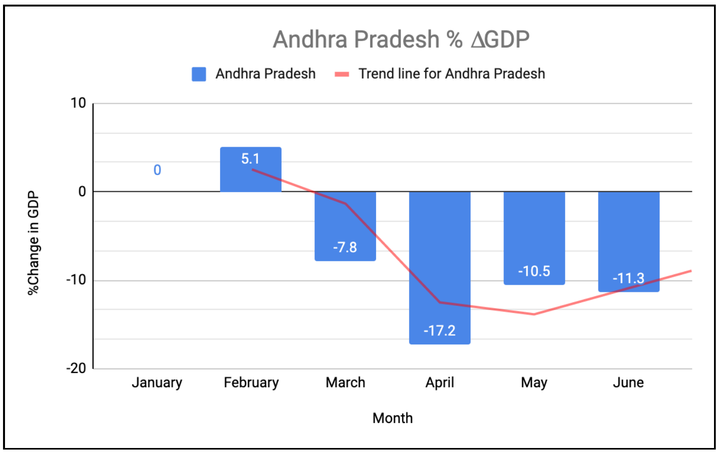
Table 8 - Summary of GDP Decline Obtained Using Framework

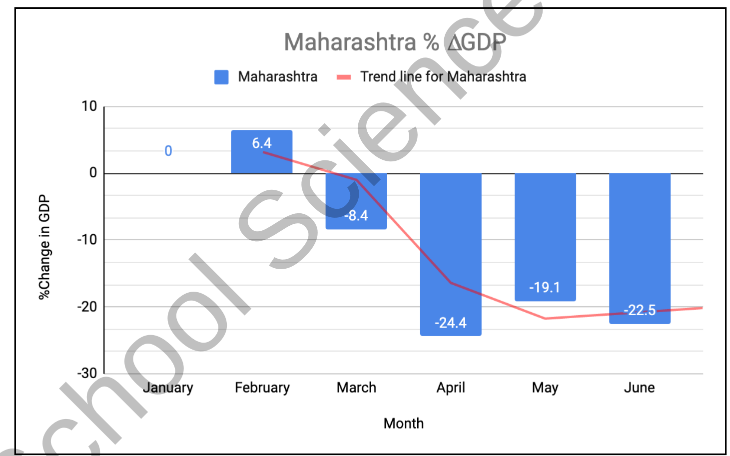


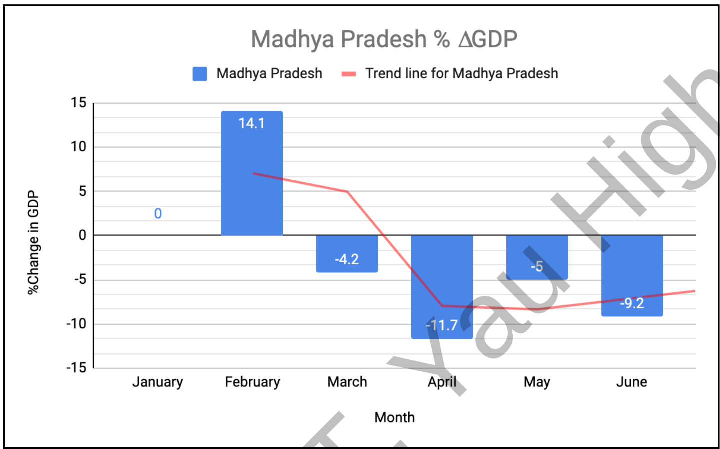


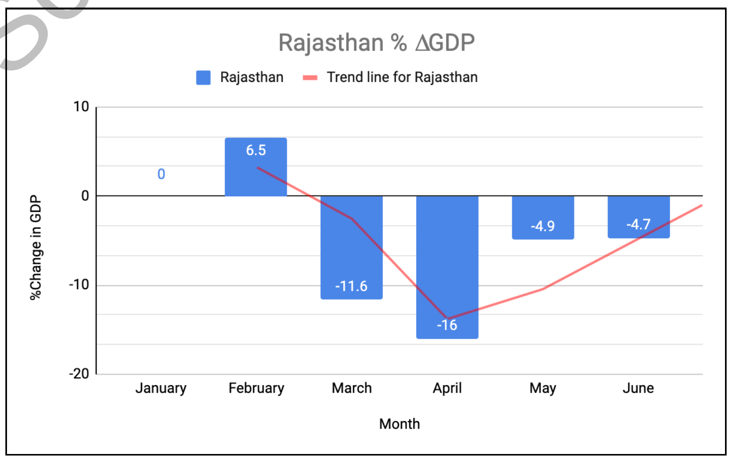












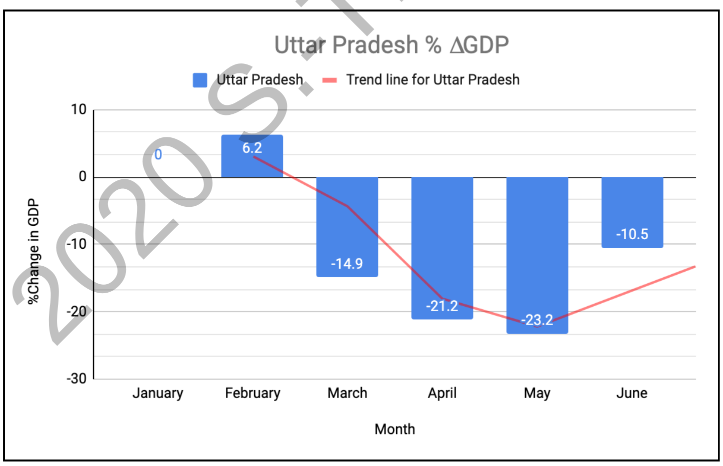


Fig 14 - Charts showing GDP Decline Obtained Using Framework

From table 8, we can observe that all states suffered a relatively small decline in March as compared to the rest of the months. This decline is expected to be small since the lockdown was effectively enforced in late March (25th March).

Industrial states, specifically Gujarat and Tamil Nadu, saw a massive decline in April of roughly 27%. This was also the month where there was a strictly enforced lockdown for the entire duration, and only primarily essential services were allowed to operate. In May, as many important industries were allowed to re-open, Gujarat and Tamil Nadu saw a recovery of 12% as compared to April. Starting June, most of the industries were allowed to operate at normal capacity, resulting in further recovery of ≈4% in June as compared to May. However, Maharashtra, another industry heavy state, saw a massive decline in April (as expected) but didn't show a recovery in May or June. We think that the reason for this could be because Maharashtra became an early epicentre for COVID cases in India, so lockdowns were enforced with further restrictions and also longer durations as compared to the nationwide lockdowns.

Rajasthan, a state whose economy depends on mining and agriculture, also followed a similar trajectory to industrial states. Given the importance of agriculture in providing essentials for the rest of the country, Rajasthan saw a sharp recovery of 11% in May as compared to April and then plateaued off in June.

Madhya Pradesh, an agricultural state, saw a small decline of 12% in April and saw a quick recovery of 7% within the next month. Hence, we can see that agriculture as an industry saw a relatively small decline in economic output due to lockdowns, as compared to other industries like mining, which reduced Rajasthan's GDP significantly.

On the other hand, West Bengal, another agricultural state, saw a decline in April of 26% but didn't see a recovery, rather a further decline of 10% in May. We hypothesize that this was due to cyclone Amphan that appeared in Mid May and caused massive devastation across the state, which is visible from the huge difference between Madhya Pradesh and West Bengal. However, by June, as the effect of Amphan subsided, West Bengal's economy also saw a recovery of 24% in June with respect to May.

Karnataka, primarily a tertiary economy and very heavy on IT-based companies, saw one of the smallest overall declines of only 13% in April and 9% in June. The likely reason for this is that many IT companies could afford to let employees work from home without compromising too much on economic output. However, we believe that Bangalore (capital of Karnataka) would have seen a higher GDP decline than the rest of the state due to counter-urbanization of migrant labourers to their hometowns.

On August 17th 2020, The State Bank of India (government owned, largest bank in India) reported an estimate of 16.5% decline for the quarter (Apr - June). The states shown in Table 8 contributed to 67% of India's GDP in the year 2018, so taking an average of the estimated decline for all of these states from Apr - June, would give us a reasonable estimate for the overall India GDP decline during this quarter. In our estimates, the average decline for India (using Electricity Consumption Data) for Apr - June was 15.7%. Thus, the difference between our estimate and the State Bank's estimate is only 0.8%, and we were able to estimate this decline much before the report came out (1.5 months before). This gives further credence to our results and demonstrates that Electricity Consumption can give accurate estimates for GDP decline in real-time.

End of Chapter 2

8. Future Work

For Night Light, we were currently forced to take only new moon dates during the data collection phase, due to the impact of moonlight on night light readings. This limited the number of readings per month to only 7 days, so we were only able to build a monthly system. Moreover, applying a cloud masking layer to remove the impact of clouds on readings further reduced the number of readings in cloudy months. However, NASA has projected to launch a cleaned version of the night light database (product code VNP46A2) by the end of 2020. This product would take into account the effect of moonlight and cloud interference. Using this, we would be able to get observations for every day of the month and could build a more accurate and even more real-time system.

For electricity consumption, the main downside is that electricity is only one type of power source. Many factories or industries operate on coal (and other fossil fuels) or renewables as well. Taking into account other energy sources like these would help in generating more accurate estimates for GDP changes. Further research could be done on using overall energy consumption for GDP estimates.

Finally, alternate data sources can also be combined to give balanced GDP estimates. Merging data sources would give more accurate estimates as different sources look at different factors to estimate GDP, so combining multiple factors would give an overall view on GDP. Further research could be done for combining Night Light and Energy Consumption for estimating GDP changes.

9. Conclusion

In conclusion, using Night Light, we were able to establish a framework to get reasonable estimates of the GDP change over a time period. Moreover, as we demonstrated in the context of COVID, the framework worked on cities across countries, both developing and developed, as well as large tier 1 cities and small tier 2 cities. Moreover, we were able to find GDP changes for time periods as small as a month, which would be further improved with NASA's new night light product.

Furthermore, using the approach of Electricity consumption, we were able to find GDP changes, updated on a monthly basis. The framework, as we demonstrated for Indian states, worked for states of all sizes. Moreover, using our framework, we were able to study the impact of the lockdown on different sectors of the economy, including primary, secondary and tertiary.

Using these alternate data sources and the framework we developed, we were able to answer many crucial questions regarding the economic impact of COVID. We were able to quantify the GDP decline in different cities and states over the months and confirm results with a balanced check using the lockdown timelines. Moreover, we were also able to quantify the effect of government policies on GDP decline, and how the varying levels of the strictness of the lockdown had different levels of impact on the economy, sector wise as well. We were also able to quantify the reduction in intercity transportation.

Our results for the impact of COVID on various states, cities, and the impact of government policies on GDP shows great promise. Such alternate data sources can be used to greatly improve quality and proactiveness of decision making by policymakers, governments and economists in the future.

References

1. *Anupam Prakash, et al. "Night-time Luminosity: Does it Brighten Understanding of Economic Activity in India?" Rbidocs.Rbi.Org.In, 2019,* https://rbidocs.rbi.org.in/rdocs/Content/PDFs/A.PDF.
2. *Beyer, Robert C. M.; Chhabra, Esha; Galdo, Virgilio; Rama, Martin. 2018. "Measuring Districts' Monthly Economic Activity from Outer Space." Policy Research Working Paper ; Paper No. 8523. World Bank, Washington, DC. © World Bank. License: CC BY 3.0 IGO."* https://openknowledge.worldbank.org/handle/10986/29996
3. "Brand India." Ministry of Commerce & Industry, Government of India, *IBEF*, 20AD, www.ibef.org/states.
4. Campo, Jacobo, and Viviana Sarmiento. "The Relationship Between Energy Consumption And GDP: Evidence From a Panel Of 10 Latin American Countries*." *Latin American Journal of Economics*, 2013, https://scielo.conicyt.cl/scielo.php?script=sci_arttext.
5. *Christopher D Elvidge, Kimberly Baugh, Mikhail Zhizhin, Feng Chi Hsu & Tilottama Ghosh (2017) VIIRS night-time lights, International Journal of Remote Sensing, 38:21, 5860-5879, DOI: 10.1080/01431161.2017.1342050*
6. Devesh Kumar, "COVID-19 India Timeline." *The Wire*, 2020, https://thewire.in/covid-19-india-timeline.
7. *Dr Miguel O Roman, et al. Black Marble User Guide Viirsland.Gsfc.Nasa.Gov, 2020,* https://viirsland.gsfc.nasa.gov/PDF/VIIRS_BlackMarbleUserGuide_V1.1.pdf.
8. *Economic data (2018 GDP PPP) for 21 cities from* https://statista.com
9. *Ferguson, Ross, et al. "Electricity Use and Economic Development." Energy Policy, Elsevier, 15 Aug. 2000,* www.sciencedirect.com/science/article/abs/pii/S0301421500000811.
10. *GDP Q2 2020 Files, Bureau of Economic Analysis U.S. Gov Department, 2020,* https://www.bea.gov/sites/default/files/2020-07/gdp2q20_adv.pdf.
11. India GSDP data from "Ministry Of Statistics And Program Implementation | Government Of India". Mospi.Nic.In, 2020, http://mospi.nic.in/data
12. Henderson, et al. "Measuring Economic Activity From Outer Space", 2012, https://edisciplinas.usp.br/pluginfile.php/4253686/mod_resource/content/1/Henderson%20et%20al%20-%20Measuring%20Economic%20Growth%20from%20Outer%20Space%20-%20AER%202012.pdf.
13. Howard Marks, "Time For Thinking", Oaktreecapital.Com, August 5th 2020, https://www.oaktreecapital.com/docs/default-source/memos/timeforthinking.pdf.
14. Joakim Wincent, et al "Monitoring The COVID-19 Crisis From Space". *MIT Sloan Management Review*, 2020, https://sloanreview.mit.edu/article/monitoring-the-covid-19-crisis-from-space/.

15. Liu, Q, et al. “Spatiotemporal Patterns of COVID-19 Impact on Human Activities and Environment in Mainland China Using Nighttime Light and Air Quality Data”. Remote Sens. 2020, 12, 1576
16. Lepara, Sabina, and Enisa Džananović. *The Relationship Between GDP and Electricity Consumption in Southeast European Countries.* ResearchGate, 2017, https://link.springer.com/chapter/10.1007%2F978-3-319-47295-9_17
17. Mohanty, Asit, and Devtosh Chaturvedi. “Relationship between Electricity Energy Consumption and GDP: Evidence from India.” *ResearchGate*, 2015, dx.doi.org/10.5539/ijef.v7n2p186.
18. Manuel, Chima. “Empirical Study Of The Relationship Between Energy Consumption And Gross Domestic Product In The USA.”, ResearchGate, 2005, www.researchgate.net/publication/267833765_Empirical_Study_Of_The_Relationship_Between_Energy_Consumption_And_Gross_Domestic_Product_In_The_USA.
19. Nallarasan, N. Power System Operation Corporation Limited,, Monthly Reports, https://posoco.in/reports/monthly-reports/
20. NOAA's Earth Observation Group, “Blue Marble Navigator - Night Lights 2012". Blue-Marble.De, 2020, https://blue-marble.de/nightlights/2012.
21. Ecowrap, *“The Good: Q1GDP Will Be Significantly Better”*, Sbi.Co.In, 2020, https://www.sbi.co.in/documents/13958/3312806/170820-SBI+Ecowrap+-+The+Good
22. Virginia Kalb, et al. "NASA VIIRS Land Products". Viirsland.Gsfc.Nasa.Gov, 2020, https://viirsland.gsfc.nasa.gov/Products/NASA/BlackMarble.html