THE TEXTILE INDUSTRY AND THE WATER QUALITY IN INDIA

Produced By: Hervin Sagnep, Naa Lamptey-Mills, Steven Lu
Summer 2020

TABLE OF CONTENT

INTRODUCTION

For our summer project, we decided to focus our analysis on the effects of the textile industry on our water quality and how it's affecting the community. In our initial brainstorming for this project, we intended on finding the correlation between the harmful business practices and its environmental impact. Our first stage of research was finding company data and documentation. However, we were able to find substantial research and documentation that led us to consider the countries that import textiles to the United States. We researched and found four datasets that will help us in the data analysis section of our project and we're excited to share our insights and findings of how the textile industry is affecting our environment.

Dataset #1:
Kaggle Search: 'India Water Quality Data'
Dataset's Name: 'IndiaAffectedWaterQualityAreas'
Short Description of Columns:
State Name: Name of State
District Name: Specific district of water source
Block Name: water sources block name
Panchayat Name: Village Council Name
Village Name: Village's Name
Habitation Name: Locality that has people or is inhabited
Quality Parameter: Type of chemical present in water
Year: Year of measurement
Number of Rows, Columns, and Missing Values:
Rows: 550243
Columns: 8
Missing Values: 0

Dataset #2:
Kaggle Search: 'India water quality data'
Dataset's Name: 'water_dataX'
Short Description of Columns:
Station Code: Station code of every listed place
Locations: address of the place where the data was collected
State: Name of state where value were measured
Temp: Average value over time
D.O. (Mg/l): Average value over time
PH: Average value over time

Conductivity (µmhos/Cm): Average value over time
B.O.D. (Mg/l): Average value over time
Nitratenann + Nitritenann (Mg/l): Average value over time
Fecal Coliform (MPN/100ml): Average value over time
Total Coliform (MPN/100ml)Mean: Average value over time
Year: 2003 - 2014

Number of Rows, Columns, and Missing Values:

Rows: 1992
Columns: 12
Missing Values:

- Station Code: 122 NAN
- Locations: 0
- State: 0
- Temp: 92 Blanks/NAN
- D.O. (Mg/l): 31 NAN
- PH: 8 NAN
- Conductivity (µmhos/Cm): 25 NAN
- B.O.D. (Mg/l): 43 Blanks/NAN
- Nitratenann + Nitritenann (Mg/l): 225 Blanks/NAN
- Fecal Coliform (MPN/100ml): 316 Blanks/NAN
- Total Coliform (MPN/100ml)Mean: 132 Blanks/NAN
- Year: 0

Dataset #3:

Acquired from: https://wits.worldbank.org/Default.aspx?lang=en
Dataset's Name: 'Indian_textile_exports.csv'
Short Description of Columns:
Year: Date of Exports
Export(US$Thousand): Amount Generated from Exports
Rows: 31
Columns: 2
Missing Values: 0

Dataset #4:

Acquired from: https://gemstat.org
Dataset's Name: 'samples.csv'
Short Description of Columns:
GEMS Station Number: State of where data was collected

Sample Date: When data was collected
Sample Time: Time of collection
Depth: How far below water levels in inches
Parameter Code: Water quality
Analysis Method Code: Form of extraction
Value Flags: NULL
Value: Measurement of unit
Unit: Volume
Data Quality: Accuracy
Rows: 237277
Columns: 10
Missing Values: NAN

Why we chose these dataset:

We chose these datasets because we decided to focus our analysis on India's water quality since it is the third largest textile contributor to the United States. We were also interested in understanding how much chemicals are present in each data sample. We wanted to determine the amount of chemicals such as iron, salinity, and other parameters are present in districts and states in India. Additionally, since India is one of the top countries that are developing rapidly among the last couple of years, we decided to focus specifically on India and what the quality of India's water is like. We hope to bring awareness and attention to the quality of water in India and how deadly these bodies of water can be.

RESEARCH STATEMENTS

Driving Question:

How is the textile industry affecting our environment?

Sub-Questions:

- Most common chemicals in each state.
- What are the costs of these effects?
- How does the textile industry affect the water quality?
- What was the average temperature each year?

Objectives:

Present our leaders with actionable results to potentially reduce environmental footprint. We hope to achieve change in terms of water consumption of textile manufacturing. We believe that the current form of manufacturing is highly damaging to our ecosystem.

DATA CLEANSING AND DOCUMENTATION PROCEDURES

Dataset #1:

Since our first dataset didn't have missing values, we decided to move forward. However, after checking the types of each variable, we noticed that the time series variables were strings instead of a date time type. Therefore, we opted to modify and convert this column from a string into a timeseries. After that, we just set the index of our dataframe to our year column. Top 5 observations of final data frame shown below.

Year	State Name	District Name	Block Name	Panchayat Name	Village Name	Habitation Name	Quality Parameter
2009-01-04	ANDHRA PRADESH	EAST GODAVARI(04)	PRATHIPADU(10)	GOKAVARAM(04)	VANTHADA(014)	VANTHADA(0404410014010400)	Salinity
2009-01-04	ANDHRA PRADESH	EAST GODAVARI(04)	PRATHIPADU(10)	GOKAVARAM(04)	PANDAVULAPALEM(022)	PANDAVULAPALEM(0404410022010400)	Fluoride
2009-01-04	ANDHRA PRADESH	EAST GODAVARI(04)	PRATHIPADU(10)	GAJJANAPUDI(06)	G. KOTHURU(023)	G. KOTHURU(0404410023010600)	Salinity
2009-01-04	ANDHRA PRADESH	EAST GODAVARI(04)	PRATHIPADU(10)	GAJJANAPUDI(06)	GAJJANAPUDI(029)	GAJJANAPUDI(0404410029010600)	Salinity
2009-01-04	ANDHRA PRADESH	EAST GODAVARI(04)	PRATHIPADU(10)	CHINTALURU(10)	CHINTALURU(028)	CHINTALURU(0404410028011000)	Salinity

Dataset #2:

The second dataset had countless missing values and blanks. To work around this we decided that we'd keep all the blanks and missing values until it is required to remove them. Eventually, when we needed to alter the missing values, we filled it with not a number string.

year	State Name	District Name	Block Name	Panchayat Name	Village Name	Habitation Name	QualityParameter	Temp	D.O. (mg/l)	PH
2014-01-01	ANDHRA PRADESH	EAST GODAVARI(04)	PRATHIPADU(10)	GOKAVARAM(04)	VANTHADA(014)	VANTHADA(0404410014010400)	Salinity	25.9	5.5	7.6
2014-01-01	ANDHRA PRADESH	EAST GODAVARI(04)	PRATHIPADU(10)	GOKAVARAM(04)	VANTHADA(014)	VANTHADA(0404410014010400)	Salinity	26.5	5.4	7.8
2014-01-01	ANDHRA PRADESH	EAST GODAVARI(04)	PRATHIPADU(10)	GOKAVARAM(04)	VANTHADA(014)	VANTHADA(0404410014010400)	Salinity	18.9	7.1	7.8
2014-01-01	ANDHRA PRADESH	EAST GODAVARI(04)	PRATHIPADU(10)	GOKAVARAM(04)	VANTHADA(014)	VANTHADA(0404410014010400)	Salinity	20.5	6.9	8
2013-01-01	ANDHRA PRADESH	EAST GODAVARI(04)	PRATHIPADU(10)	GOKAVARAM(04)	VANTHADA(014)	VANTHADA(0404410014010400)	Salinity	27	6	7.3

CONDUCTIVITY (µmhos/cm)	B.O.D. (mg/l)	NITRATENANN+ NITRITENANN (mg/l)	FECAL COLIFORM (MPN/100ml)	TOTAL COLIFORM (MPN/100ml)Mean
569	1.3	0.9	18	141
310	0.8	2.4	7	107
516	0.7	0.5	2	1238
505	0.9	0.7	2	1563
432	1.2	3	3	78

Since both datasets had a common variable, states. We decided that this will be the best variable to merge the data set by. We believe that merging the datasets will provide further information about districts, and the different metrics in each specific district and block name. We also used the year variable as the index instead of traditional indexing. After merging the data set, we decided to drop the unnecessary variables to make our dataframe more presentable.

In addition to merging both datasets, we kept the individual datasets as well for individual analysis if required.

Dataset #3:

For our third dataset, the cleansing part of it consisted of separating and expanding the dataset from one column to ten different columns. We simply used the split function to manipulate the data set and ended up with the following:

GEMS Station Number	Sample Date	Sample Time	Depth	Parameter Code	Analysis Method Code	Value Flags	Value	Unit	Data Quality
IND00001	1991-04-04	12:00	0.6	H-T	T-COL-EDTA-EBT		116.0	mg/l	Fair
IND00001	1991-04-04	12:00	0.6	TP	COL-SnCl-SA-PPS		0.0	mg/l	Poor
IND00001	1991-05-07	12:00	0.6	TKN	TKN-T-COL		3.36	mg/l	Fair
IND00001	1991-05-07	12:00	0.6	NH3N	COL-NES		1.12	mg/l	Fair
IND00001	1991-05-07	12:00	0.6	O2-Dis	T-COL-I-AZD		8.2	mg/l	Fair

Dataset #4:

For the fourth dataset, we needed to change the index of the dataframe to a time data type so we are able to create a proper time series. We also needed to remove the commas in the first column so matplotlib would accept the values.

Year	Export(US$Thousand)
1988-01-01	3148861.47
1989-01-01	4080891.82
1990-01-01	4899971.21
1991-01-01	4882667.33
1992-01-01	5707423.89

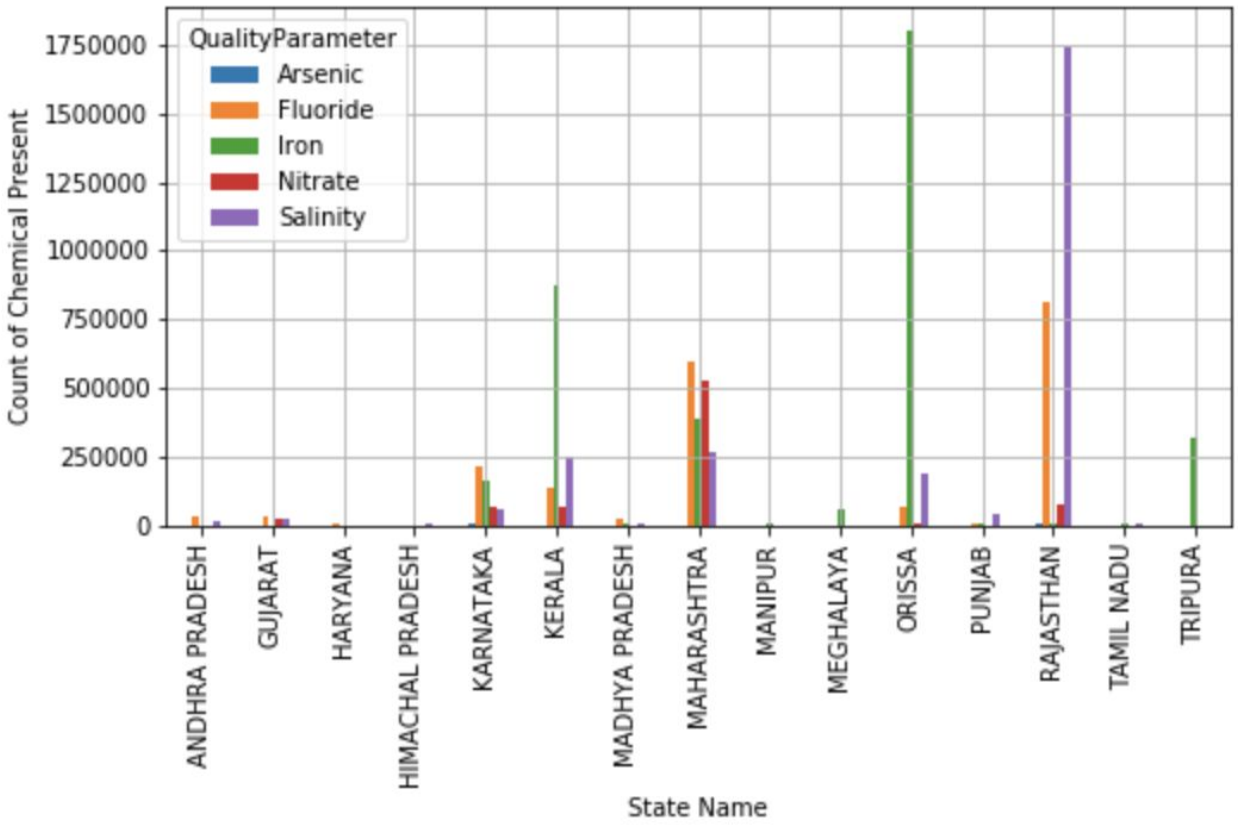
Our project consisted of utilizing up to four data frames to analyze the water quality in India and how textile imports have devastated the environment.

DATA ANALYSIS

Sub Question 1: Most common chemicals in each state.

For our first question, we decided to look at the chemicals most present in each state. With this information, we'd be able to understand the potential effects of the water and understand the potential hazard behind water sources with larger volumes of specific chemicals. From the figure below, we can see that Orissa is most polluted with iron and Rajasthan is most polluted with Salinity.

Some potential negative side effects of ingesting excessive amounts of iron includes damage to healthy skin cells, mutation in the genes, plumbing issues due to excessive residue left behind, and lack of quality in food and beverages. This is an important factor to consider not only due to the factors presented previously, but also the potential long-term effects to the communities.

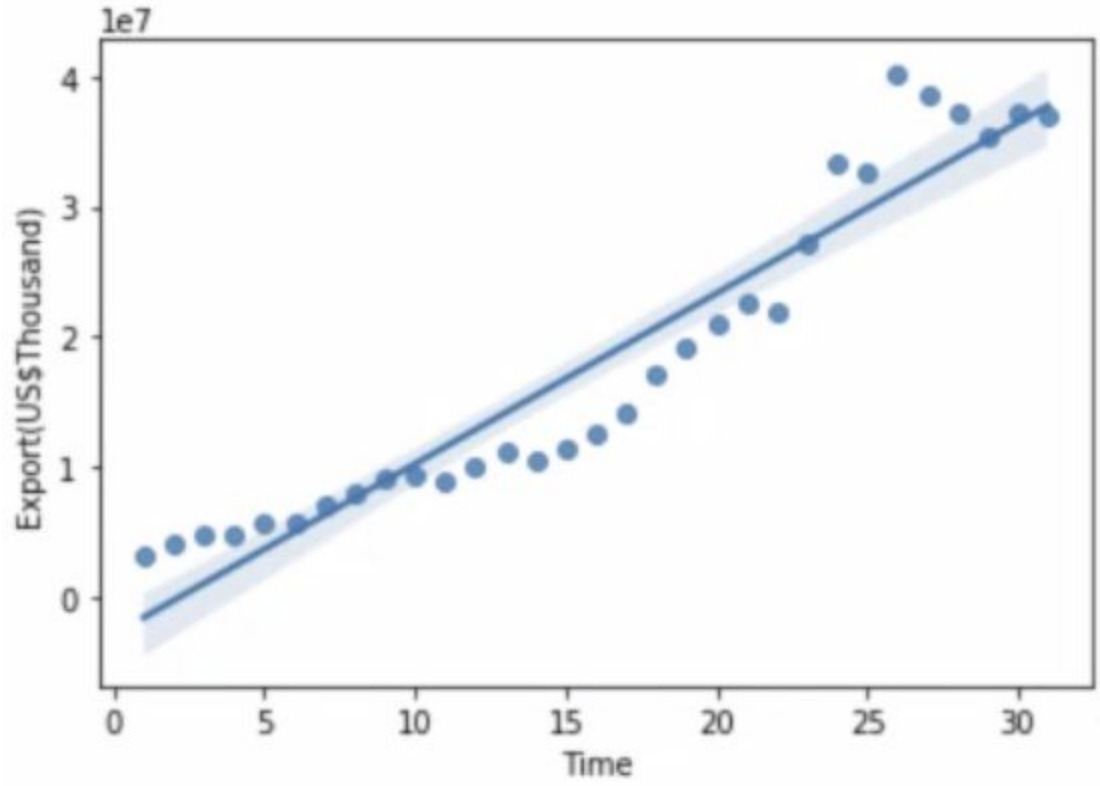


Sub Question 2: What are the costs of these effects?

As India is still a developing country with over 1.3 billion residents, the quality of water in India plays a major role in the lifestyle of these residents. Almost 80% of India's surface water

is contaminated. With millions of litres of wastewater being poured into bodies of water daily, it will create an upstream of lower economic growth for communities in the downstream areas, which will eventually reduce GDP growth in these communities by up to ⅓.

Sub Question 3: How does the textile industry affect the water quality?

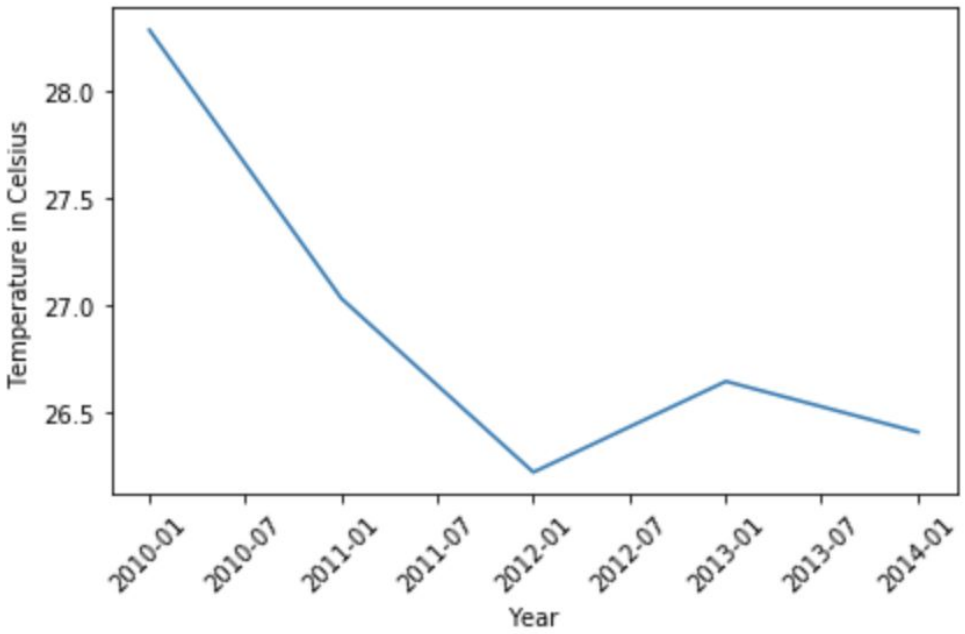


The textile industry is one of the most rapidly growing industries in India. As seen above, you can see the value exports of Indian has a linear upward trend over time which points to the fact that the quality of water will decrease as the exports of India increases. The uses of various chemicals and dyes while manufacturing fast-fashion textiles generates an unprecedented quantity of chemically polluted waters and sludge.

Chemically polluted wastewater will result in downgraded quality of soil. Especially when mixing these natural resources with its dependent environments. Textile companies are now facing major problems in environmental sustainability, as their manufacturing results in consequences such as high quantities of solid and liquid waste.

As previously mentioned, with millions of litres of contaminated water being poured into bodies of water weekly, it will create an increase in difficulties growing the country GDP due to the lack of consideration for citizens in the downstream areas.

Sub Question 4: What was the average temperature each year?



As seen above, 2004 water temperature was at an all time high of 85.1F (28.29C) with recent years slowly decreasing. We believe that this is due to specific policies passed in recent years to maintain and control the amount of wastewater poured into sewages.

The rising temperature of water will affect the aquatic organisms in a magnitude of ways. Most aquatic organisms are adaptable to a narrow range of water temperature. If the temperature increases or decreases too rapidly, it will result in immediate death of large populations of aquatic organisms which can eventually lead to a shortage in India's seafood supply chain.

CONCLUSION

Analysis Conclusion:

Our team intention with this project was to create insightful information and statistics about how deadly the textile industry is on certain communities and we decided to focus on India as our dominant country. We decide to analyze datasets about the water quality and temperature levels to see what exactly the water quality is like and how it can potentially affect the community surrounded by these bodies of water.

We decided to analyze the chemicals present in different states to understand the most prominent chemicals and research the potential side effects of these chemicals. We've learned that excessive amounts of iron may lead to damaged skin cells, mutation of the genes, and plumbing issues due to heavy residue left behind.

In addition, we studied the side effects of pouring waste water into bodies which can lead to destructive economic growth for communities in the downstream areas. With economic inequality for members of the downstream areas, it will eventually lead to a reduction in GDP growth in these communities by up to 33%.

We then focused on study on the amount of textile exported out of India and we have discovered an upward linear trend. We can confidently say that the exports will continue to rise which is correlated to the amount of wastewater being dumped into bodies of water. This will eventually result in downgraded quality of soil and difficulties growing vegetation in downstream areas.

Lastly, we wanted to focus on the temperature levels of water in different areas of India and we do notice a shockingly high value in certain years. The temperature of water is crucial to environmental sustainability because communities that rely heavily on seafood. Increase of temperature to sea levels will lead to immediate deaths of many aquatic organisms that have difficulties adapting to changing temperatures.

Our analysis solidified our hypothesis of how the textile industry is impacting our environment and ecosystem. We were able to look at the chemicals present in water, its effects, and the temperature of bodies of water in different areas of India.

Project Conclusion:

As a small team, we decided to focus on applying what we've learned throughout the program and see how data can be used to potentially influence decision makers.

Our group’s data analysis was a small contribution to existing fashion and sustainability research. While we were able to analyze textiles' negative effects on the water quality in various states in India. We were able to learn how public health and economics were impacted by the textile industry. Our main contribution is demonstrating that there is a correlation between textile waste and the environment. We would have included data from other countries and corporations to create more clear calls to action and insights. We hope that this project will serve as motivation for other researchers and a re-consideration of what sustainability is in the textile industry.