



# FTA & ECPAT HT Project



# Meet The Team



## ECPAT International



**Tyler Hayek**  
B.S. Information Tech  
MBS (Cybersecurity)

**Hernán Chacón**  
Mining Engineer  
MBS (Data Analytics)

**Myra Shama**  
B.S. in Electrical  
Engineering



**Francisco Almachi-Vera**  
B.S. in Mechanical  
Engineering



**Albert Joseph**  
B.S. in Industrial and  
Systems Engineering



**Willio Jeanvilma**  
B.S. Computer  
Science SASN

## RUTGERS UNIVERSITY ADVISOR

**Christie Nelson, Ph.D. in Operations  
Research**  
Assistant Research Professor

## PROGRAM MENTORS

**John Betak, Ph.D in Geography and  
Artificial Intelligence.**  
Company Mentor



## Background

- \* Human trafficking is the second most committed criminal activity in the world.
- \* Sex trafficking continues to be an issue for minors in the US and trends have increased since the start of the pandemic.

## Problem Statement

Human trafficking is becoming harder to detect and catch with the use of social media and other websites that allow users to become anonymous.



## Methods

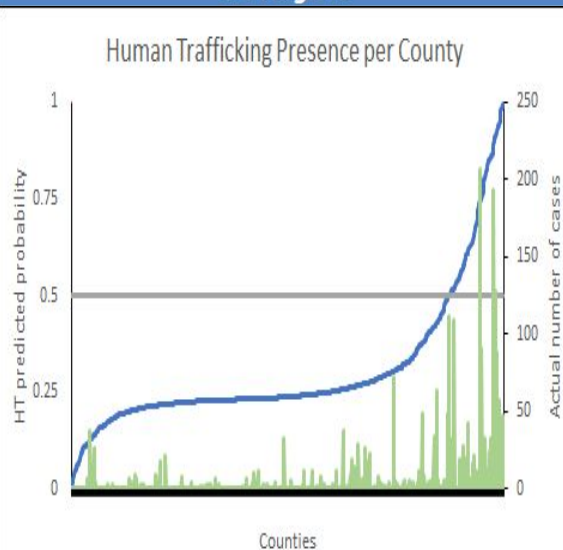


- Python: used to clean, manipulate, and acquire data.
- JMP: used for logistic regression within socioeconomic data.



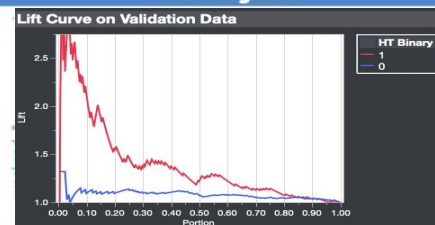
Twitter: compiled tweets that contained specific keywords, & hashtags from specified locations in a given date range.

## Analysis



We fitted a logistic regression and received a score of 75%

## Analysis



Above is the Lift curve for the validation dataset results

## Discussion & Conclusion

- \* COVID-19 has caused a dip in Tweets advertising illegal escort services
- \* Major cities and tourist/travel locations have larger volume of tweets referencing commercial sex acts.

## Future Work

- \* Apply similar methodology to other social media platforms and websites where traffickers are known to be more prevalent.
- \* Analyze live streaming or gaming chats for presence of traffickers or grooming.





# **1: Background on the** **... Company & Industry**

- **ECPAT**

- Founded in 1990
- Global organization
- Directed to end the sexual exploitation of children through the means of prostitution, trafficking, online, travel/tourism sector

- **FTA (Federal Transit Administration)**

- Founded in 1964
- Provides financial and technical assistance to local public transit systems
- Oversees safety measures



## 2: Problem Statement / Goals

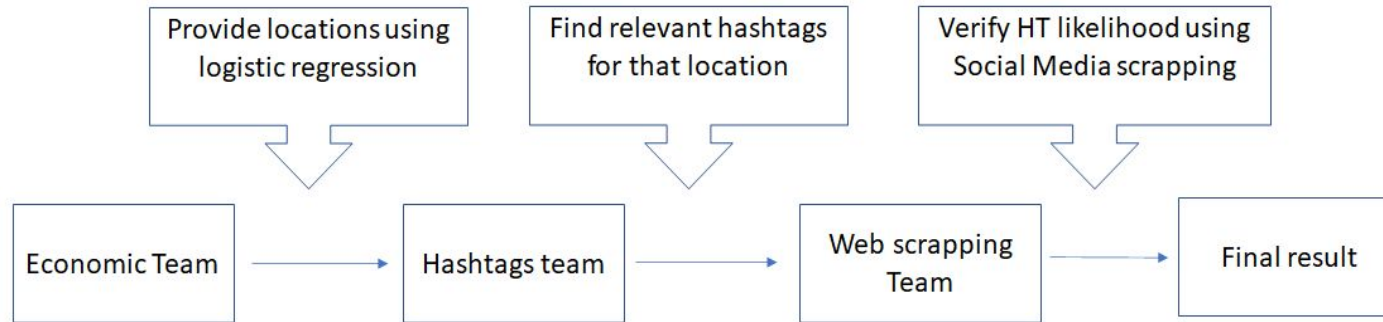
### Problem

- Human trafficking is becoming harder to detect and catch with the use of social media and other websites that allow users to become anonymous.

### Goals

- The goal of our team project is to deliver a systematic process to predict human trafficking by looking at socioeconomic data and twitter hashtags.
- Analyze keywords and indicators corresponding to human trafficking in order to find trends.

Goal: Systematic process to confirm if there's human trafficking in a given location



# 3: Project Approach/Methods

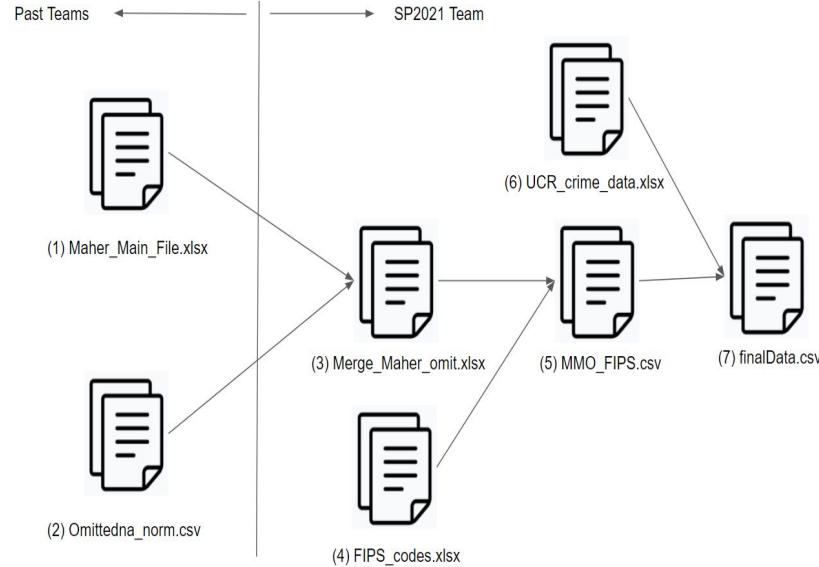


## Data Collection

Previous teams: Excel data extraction from Census, BLS, NIBIRS, and UCR.

## Data Curation

Python code to integrate data and obtain preliminary results.



## Analysis

Use of logistic regression to complement models and validate results.

## Visualization

Visualization will be done using Tableau, JMP, and Excel.

# 4: Project Timeline

## Project Planning and assessment of current information

Review techniques and findings from previous group and plan out strategies with Mentor.

### Milestone 2



## Implementation

Conduct social media website scraping and socioeconomic data analysis.

### Milestone 4

## Model Review

Use Socioeconomic economic data to predict HT hotspot and verify hotspot with social media scraping.

### Milestone 1



## Discovery and Research

Filter out the necessary websites and research existing data; discover where sex trafficking is likely to occur.

### Milestone 3



## Data Analysis

Create logistic regression with socioeconomic data. Use predictors to find hashtags on twitter.

### Milestone 5



# 5: Results

We fitted a logistic regression using finalData.csv  
 + We obtained a probability for HT presence and absence

Effect Summary	
Source	
Unemployed	
Births	
actual_robbery_unarmed	
actual_burg_force_entry	
actual_robbery_with_a_gun	
actual_arson_storage	
actual_assault_unarmed	
actual_rape_attempted	
Poverty	
DomesticMigration	

Confusion Matrix		
Training		
Actual	Predicted	
HT Binary	1	0
1	72	197
0	36	575
Validation		
Actual	Predicted	
HT Binary	1	0
1	20	51
0	21	201

Score: **75.43%**

Year	County	Prob[1]	Prob[0]	Most Likely HT Binary	HT Binary	Correct ?
2018	Maricopa, AZ	0.911	0.089	1	1	Yes
2018	Howard, MD	0.190	0.810	0	0	Yes
2018	Bexar, TX	0.983	0.017	1	1	Yes
2018	Montgomery, MD	0.771	0.229	1	0	No
2018	Douglas, CO	0.247	0.753	0	0	Yes
2018	Hutchinson, TX	0.240	0.760	0	0	Yes

Some of the hashtags we found in our predicted locations:

#bookme  
 #flyme2you  
 #Date[city]  
 #Adultwork  
 #luxurycompanion



# 6: Recommendations



## Student Team

- Learned the different components of web scraping.
- Learned to work with different datasets.



## Future Methodology

- Other social media scrapers (Tik Tok and Instagram) and other enterprise level scrapers.



## FTA/ECPAT

- Gained new potential tools and keywords to find locations with Human trafficking.



## Target Victims

- Traffickers look more for emotionally vulnerable victims whom they can more easily manipulate.
- This could be incorporated to future models.

## ISSUED FOR

John Betak, Ph.D.

## RUTGERS ADVISORS

Christie Nelson, Ph.D.

## GET IN TOUCH

Francisco Almachi-Vera  
fa382@scarletmail.rutgers.edu

Albert Joseph  
abj37@rutgers.edu

Myra Shama  
ms2913@rutgers.edu,

Tyler Hayek  
tvh21@scarletmail.rutgers.edu

Willio Jeanvilma  
wj150@rutgers.edu

Hernán Chacón  
hfc15@scarletmail.rutgers.edu



# ECPAT/FTA Human Trafficking Social Media Analysis

Students: Francisco Almachi-Vera (Lead), Myra Shama, Tyler Hayek, Willio Jeanvilma, Albert Joseph, Hernán Chacon

Company Mentor: Dr. John Betak | MBS Mentor: Dr. Christie Nelson



## Background

- \* Human trafficking is the second most committed criminal activity in the world.
- \* Sex trafficking continues to be an issue for minors in the US and trends have increased since the start of the pandemic.

## Problem Statement

Human trafficking is becoming harder to detect and catch with the use of social media and other websites that allow users to become anonymous.



## Methods

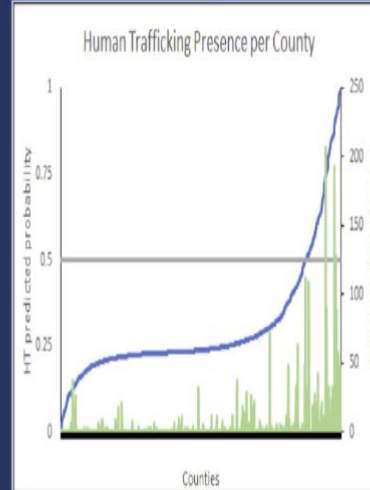


- Python: used to clean, manipulate, and acquire data.
- JMP: used for logistic regression within socioeconomic data.



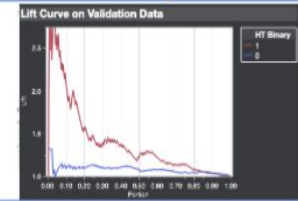
Twitter: compiled tweets that contained specific keywords, & hashtags from specified locations in a given date range.

## Analysis



We fitted a logistic regression and received a score of 75%

## Analysis



Above is the Lift curve for the validation dataset results

## Discussion & Conclusion

- \* COVID-19 has caused a dip in Tweets advertising illegal escort services
- \* Major cities and tourist travel locations have larger volume of tweets referencing commercial sex acts.

## Future Work

- \* Apply similar methodology to other social media platforms and websites where traffickers are known to be more prevalent.
- \* Analyze live streaming or gaming chats for presence of traffickers or grooming.



# THANK YOU

