11.154/11.454
Big Data, Visualization, and Society
Fall 2021 - Central American Migration

Monday + Wednesday : 9:00-10:30
Location Room : 9-451

Instructors:
Sarah Williams
Department of Urban Studies and Planning
E14-140 | sew@mit.edu |
Office Hours : https://calendly.com/sew-williams/meet

Ashley Louie:
Department of Urban Studies and Planning
E14-140 | alouie@mit.edu
Office Hours: Wednesday, Thursday 1-3pm via zoom
https://mit.zoom.us/my/a.louie

TA: Asher Simon
(asimon@mit.edu)
Office Hours: Wednesday, 10:30am-12pm
TA: Nikolas McGlashan
Office Hours:
Description
Big data has become a major part of how we understand the city, make decisions, and project what might happen next. This class studies data visualization as a way for architects, planners and policy experts to communicate with the public. Through this class, students will develop technical skills to work with big data to answer or expose urban issues, which include cleaning and aggregating data in python, D3, and other web-based visualization software.

Objectives
By the completion of the course students will:

- Grasp the conceptual and technical processes behind data-driven narratives.
- Learn how to tell stories with data.
- Analyze data using Python, Pandas, and other python based libraries.
- Visualize data using visualization tools including D3
- Learn how to use data visualization for policy change.
- Learn how to develop a website and use CSS and Javascript.

Focus: Central American Migration

Each year the class works with a “client” focusing on a theme, this year’s theme was developed in coordination with the World Food Program and the Migration Policy Institute to look at the root causes of migration. Central American migration to the United States has shaped regional flows for decades, but recent changes in the volume and composition of these flows have drawn unprecedented attention from regional governments seeking to reduce irregular migration. Combined encounters of migrants from Guatemala, El Salvador, and Honduras at the U.S.-Mexico border reached 1.6 million since 2017, surpassing the number of encounters of Mexican migrants in four of the last five fiscal years. Since 2018, furthermore, the region has witnessed larger and more frequent mass migration movements, including significant shares of unaccompanied minors and families, primarily bound for the United States and seeking employment opportunities, family reunification, and humanitarian protection.

Though no one single factor drives migration on its own, economic stagnation in Central America has been a persistent trigger that has worsened because of the COVID-19 pandemic. Each year, the number of young people entering the labor market surpasses the number of jobs available, and many who do not find jobs decide to migrate instead. In 2020, the economic pressures were even more extreme, as GDP contracted by 2 percent in Guatemala, 8 percent in El Salvador, and 9 percent in Honduras. That same year, projections from the UN Economic Commission for Latin America and the Caribbean (ECLAC) suggested that more than half of Guatemalans and Hondurans and 40 percent of Salvadorans lived in poverty.
The effects of the pandemic and poverty have also magnified levels of food insecurity in the region. According to the UN World Food Programme (WFP), the number of Guatemalans, Salvadorans, and Hondurans affected by moderate or severe food insecurity nearly quadrupled from 4.8 million in 2019 to 17.3 million by fall 2020. The dramatic increase was associated with an increase in reported intentions to migrate from 8 to 15 percent over the same period.\(^6\)

Violence, crime, and corruption are also key drivers of migration. Though they appear to be falling, homicide rates in El Salvador and Honduras continue to be among the world’s highest, and parts of Guatemala are equally violent.\(^7\) About one in five residents across these countries report being victims of crimes every year. And nearly one-tenth of Hondurans and Salvadorans experience extortion annually, paying gangs and local crime groups just so they can live in their homes or run small businesses.\(^8\) Additionally, high-level governmental corruption as well as lower-level corruption amongst security actors and public officials can also undermine people’s faith in institutions and factor amongst drivers of out-migration.

Less present in ongoing regional policy dialogues compared to other factors, but equally as important are the worsening impacts of climate-related shocks, both in terms of slow onset and sudden onset hazards. These events, compounded by environments of high vulnerability, violence, unemployment, and limited access to social protection schemes, overwhelm community resilience. Additionally, Guatemala, El Salvador, and Honduras have borne the brunt of intense climate shocks such as Hurricanes Eta and Iota in November of 2020, while
local agricultural markets were decimated in 2018 by one of the worst droughts in the last 40 years.⁹

At the same time, there is renewed political interest from Canada to Panama to collaboratively address the root causes of migration and displacement. These efforts have followed advancements in regional frameworks that seek to promote safe, orderly, and regular migration with the goal of benefiting migrants and origin and destination communities alike, while contributing to sustainable development. A fundamental example is the Global Compact for Safe, Orderly and Regular Migration which underscores the importance of addressing the adverse drivers and structural factors that compel people to leave their country of origin and calls for establishing safe pathways for persons affected by disasters, environmental degradation and climate change. At a regional level, the Central American Integration System (SICA) adopted the Policy Proposal for Comprehensive Regional Migration in 2018 to ensure that intra-regional migration flows are governed by the principles of human rights and security.

Governments in the region thus face a unique and timely opportunity to shift from an enforcement-centered strategy to a comprehensive migration management system by leveraging renewed interest in cooperation. The U.S. administration’s recent proposal to devise a “collaborative migration management strategy” in the region may be the most concrete and influential example, but is it not the only one.¹⁰ The governments of Mexico, Guatemala, El Salvador, and Honduras have previously called for the region to address migration based on a principle of co-responsibility.¹¹ More recently, Canada’s Immigration Ministry expressed interest in increasing its capacity to resettle more Central American refugees.¹² And other examples include a request by Panama’s Foreign Minister for regional cooperation and responsibility to address migration flows from Central America and Venezuela.¹³

Data:
Students will be able to use data collected from approximately 5,000 household interviews and over 6,000 web survey responses across the three countries conducted as part of a joint initiative between WFP and the International Organization for Migration (IOM) to better understand these factors that cause migration. The household interviews included questions about personal finances, food security, remittances, among questions revolving around the motivations to migrate. Students will be encouraged to use this rare data set and encouraged to combine it with other data sets found on the web to help support their stories.

Data Ethics:
It is important for us to recognize that the data you are using represents people and should be recognized as such. While the data is anonymized, which means you can not trace an individual in the study, careful consideration should be given toward protecting the people described in the study. Please make sure you ask yourself whether the analysis you create from the data can do them any harm. Please also reflect on your own positionality within the analysis you perform with the data. Also this data used in this class should not be shared with others without the express consent of Sarah Williams.
Data:
All data resources are provided in the Google Drive Data Resources folder including metadata about the files. We also include links to online data portals which can be helpful for finding external data sets.

Readings:


Schwabish, Jonathan and Alice Feng. “Do No Harm Guide: Applying Equity Awareness in Data Visualization” Urban Institute, June 9th, 2021


Migration Literature Review
(You will be asked to pick one article to read from this literature review)

Class Schedule:

<table>
<thead>
<tr>
<th>Monday</th>
<th>Wednesday</th>
<th>Assignments Due</th>
<th>Extras</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ashley - Sep 8 Thursday at midnight</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PSET0 due Sunday (9/12) at midnight: Setting up your Environment and Intro to Python Tutorial (T01)</td>
<td></td>
</tr>
<tr>
<td>Sep 13</td>
<td>Sep 15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 2</td>
<td>Week 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T02: Python - Logic, Lists &amp;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T03: Intermediate Python - Loops &amp;</td>
<td></td>
</tr>
<tr>
<td>Week 3</td>
<td>Sep 20</td>
<td>Sep 22</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>T04: Package and NumPy / Pandas 1</td>
<td>Explain the Class Topics / Data Visualization Examples</td>
<td>PSET1 Due</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Speakers: World Food Program and Migration Policy Institute</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1) &quot;The Real Root Causes of America's Border Crisis&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2) Watch this vimeo -- Design Homeless <a href="https://vimeo.com/30391522">https://vimeo.com/30391522</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3) 3121 desperate journeys exposing a week of Chaos under Trump's zero tolerance. (Problem Set 3 Release)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Week 4</th>
<th>Sep 27</th>
<th>Sep 29</th>
</tr>
</thead>
<tbody>
<tr>
<td>T05: Pandas 2 + MatPlotLib (A3 release)</td>
<td>T06: Analyzing Migration Study Data</td>
<td>PSET2 Due</td>
</tr>
<tr>
<td>Read:</td>
<td>Read:</td>
<td></td>
</tr>
<tr>
<td>Top 6 Python Libraries for Visualization: Which one to Use?</td>
<td>1) One article of your choice from the Literature Review. I will ask you at the beginning of class to give me a two min summary of one of the articles.</td>
<td></td>
</tr>
<tr>
<td>2) The Great Climate Migration</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Week 4* must meet with TA or instructor about project ideas.
<table>
<thead>
<tr>
<th>Date</th>
<th>Activity</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oct 4</td>
<td>Ashley</td>
<td>- Week 5 Project Ideas - Science Fair - Student Presentations</td>
</tr>
<tr>
<td>Oct 6</td>
<td>T07: Examples of Midterms / Bokeh Visualization Tools</td>
<td>- Read: Better Data Visualization (pages 1-52 - don’t worry its lots of images. )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 2) Chapter 5 of “Do no Harm Guide.” Data behind the Viz.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- (Problem Set 4 Release) Science Fair &amp; PSET4 Due</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- maybe additional viz session</td>
</tr>
<tr>
<td>Oct 11</td>
<td>Storytelling with Data and Developing Wireframes</td>
<td>- Watch - Design with Data</td>
</tr>
<tr>
<td>Oct 13</td>
<td></td>
<td>- PSET5 Due</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- meet with TA or instructor about project ideas</td>
</tr>
<tr>
<td>Oct 18</td>
<td>Introduction to HTML / CSS/ Website</td>
<td>- Scrolly Telling / Integrating Multimedia in Website</td>
</tr>
<tr>
<td>Oct 20</td>
<td></td>
<td>- Work on Midterm</td>
</tr>
<tr>
<td></td>
<td>Ashley</td>
<td>- Week 7 Introduction to Javascript - Introduction to D3</td>
</tr>
<tr>
<td></td>
<td>MIDTERM</td>
<td>- (Problem Set 7 Release)</td>
</tr>
<tr>
<td></td>
<td>(A6 release)</td>
<td>- Midterm Submission</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- optional Github tutorial</td>
</tr>
<tr>
<td>Nov 1</td>
<td>Introduction to Javascript</td>
<td>- PSET6 Due: Website - Scrolly telling - revised storyboard</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- must meet with TA or instructor about project ideas</td>
</tr>
<tr>
<td>Nov 8</td>
<td>Nov 10</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>D3: Working with External Data Source</td>
<td>D3: Scales and Axis (Problem Set 8 Release)</td>
<td>PSET 7 Due</td>
</tr>
<tr>
<td>Ashley - Nov 15</td>
<td>Nov 17</td>
<td></td>
</tr>
<tr>
<td>Week 11</td>
<td>D3: Making it Dynamic</td>
<td>D3: Making it Interactive (Problem Set 9 Release)</td>
</tr>
<tr>
<td>Nov 22</td>
<td>Nov 24</td>
<td></td>
</tr>
<tr>
<td>Week 12</td>
<td>D3: Maps and Networks</td>
<td>D3: Hierarchical Data</td>
</tr>
<tr>
<td>Nov 29</td>
<td>Dec 1</td>
<td></td>
</tr>
<tr>
<td>Week 13</td>
<td>Draft Review</td>
<td>Draft Review</td>
</tr>
<tr>
<td>Ashley - Dec 6</td>
<td>Ashley - Dec 8</td>
<td></td>
</tr>
<tr>
<td>Week 14</td>
<td>Final Projects due Review</td>
<td>Final Projects due Review</td>
</tr>
</tbody>
</table>

**Grading + Assignments**
Final grades will be calculated according to the following:

**50% Technical Assignments**
Of the first 9 problem sets we will drop the lowest grade - this means if you do not submit one problem set it will be dropped. If you decide to do all the problem sets you will get extra credit.

**15% Midterm**
(includes science fair assignment)

**30% Final Project**
(includes story board review, draft presentations which will be done in class) The final project will be an interactive visualization analyzing data which hopes to help set a platform for the next mayoral election. Students will be working towards the final project the whole semester. Your midterm is a pitch for the final project.
5% class participation (this includes making sure you meet the requirements for the extra meetings or workshops)

RESOURCES:

Resources for Programming:
Python Crash Course (https://github.com/ehmatthes/pcc_2e/)
Automate Boring Stuff With Python (https://automatetheboringstuff.com/)
D3 - Scott Murray's "Interactive Data Visualization for the Web" (2nd edition)
D3 - https://wattenberger.com/blog/d3#intro
Introduction to Jupyter Notebooks
The programming historian (novice friend, peer reviewed tutorials)

Recent Online Data Visualization Classes
https://infovis.fh-potsdam.de/tutorials/
https://datajournalism.com/

Popular data visualization software:
Visualizing Data Listing of Resources
Review of 6 Python visualization libraries
Datawrapper - Easy to use online software to create visualizations from prepared data. Used by many news sites and online publications.
Google spreadsheet charts - Visualization tools inside spreadsheet software.
Excel charts Visualization tools inside spreadsheet software.
Tableau Often considered the standard for professional data visualization. Also best for fast data exploration. You can register on their site with your student info, and they will give you Tableau Desktop for free - it’s more powerful and easier to use than Tableau Public. See for example -- https://public.tableau.com/en-us/s/covid-19-viz-gallery
d3 - Best software for best looking online interactive visualizations; requiring writing code in Javascript. Covered in advanced visualization courses offered by the MS in Data Analysis and Data Visualization program at the GC.
Observablehq - Working with and visualizing data using d3 inside an interactive notebook (developed by the same person who made d3).

Visualization Inspiration:
Flowing Data
Information Is Beautiful Awards
Information Is Beautiful Awards Showcase
The Pudding
Visualizing Data
New York Times Year in Graphics 2020
New York Times a Year in Graphics 2019
New York Times a Year in Graphics 2018
New York Times a Year in Graphics 2017

Migrant Data Visualizations
Stories Behind a Line
What would you do? Take an immigrant’s journey.
Migration Trail
The Missing Migrant Map
El costo de cruzar la frontera de contrabando
Bussed Out: How Americans Move Their Homeless

Wireframe and Storyboarding Tutorial:
How to Create your first Wireframe -
The Ultimate Guide to Wireframe Design
Storyboard Your Project
Structures of Stories

Videos Explaining How Popular Data Visualizations Were Made;
Nadeieh Bremer (Homelessness)
Pedro Cruz (Immigration to the US)

Podcasts:
Cool Infographics has a number of really cool podcasts on data viz. One of my favorites on this list is the policyviz podcast.

How do designers use data visualization to take the numb out of numbers?

TED and TEDx talks by influential data visualization designers, artists, and scientists
1 - Giorgia Lupi
2 - Manuel Lima
3 - Hans Rosling
4 - Aaron Koblin
5 - Jer Thorp
6 - David McCandless

Return to in-person teaching/learning
MIT and DUSP are excited at the opportunity for a return to in-person teaching and learning after 2+ semesters of remote life. To ensure that all classes can and will be delivered in person, MIT has worked hard to put into place policies, procedures and technologies to maximize the likelihood of a safe and uninterrupted semester. That said, the ever-evolving pandemic means we need to be prepared. If any student in class tests positive for covid-19, MIT Medical has
established clear procedures for ensuring safety of everyone and MIT’s Class Notification and Support Team will help that student continue learning with the least possible disruption. If any of the Instructors in this class are unable to attend in person due to covid-19, we will work together with the leadership of DUSP, the School, and the Institute to ensure minimum disruption.

In the case of the need for any remote teaching, we will use this “zoom classroom”:

Join URL: https://mit.zoom.us/j/97897130940

**Slack Course Chat Room:**
Let’s try to help each other! Let’s use Slack to post questions and receive help from your fellow students. Slack workspace URL is: mit-11-454.slack.com We will invite all registered students via Email.

**Email Etiquette**
Please give us at least 24 hours to reply to your emails, and we will do the same for you. Please put the course number in the Subject line and remember to sign your email with your name. We expect the language and structure of your emails to be professional. This includes punctuation, salutations/signature, etc.

**Statement about Race:**
The topic of this class deals with migration patterns from Central America. The roots of migration stem from climate change, extreme poverty, violence and unstable governments. The complex issues that cause migration can be linked to a history of colonization and attempts to decolonize. Inherently linked are issues of racism toward the indiginous people of Central America combined with the racist behavior towards the migrants once they arrive in the United States. It is important to remember these relationships as we work on this extremely sensitive topic, this is why we are embedding the Migration Policy Institute and the World Food Program into our work to help provide context to the long-standing issues of the migrants.

**Land Acknowledgement Statement:**
MIT acknowledges Indigenous Peoples as the traditional stewards of the land, and the enduring relationship that exists between them and their traditional territories. The land on which we sit is the traditional unceded territory of the Wampanoag Nation. We acknowledge the painful history of genocide and forced occupation of their territory, and we honor and respect the many diverse indigenous people connected to this land on which we gather from time immemorial.

**Inclusive Classroom**
MIT values an inclusive environment. I hope to foster a sense of community in this classroom and consider this classroom to be a place where you will be treated with respect. I welcome individuals of all backgrounds, beliefs, ethnicities, national origins, gender identities, sexual
orientations, religious and political affiliations – and other visible and nonvisible differences. All members of this class are expected to contribute to a respectful, welcoming, and inclusive environment for every other member of the class. If this standard is not being upheld, please feel free to speak with me.

**Special Accommodations:**
MIT is committed to the principle of equal access. Students who need disability accommodations are encouraged to speak with Disability and Access Services (DAS), prior to or early in the semester so that accommodation requests can be evaluated and addressed in a timely fashion. If you have a disability and are not planning to use accommodations, it is still recommended that you meet with DAS staff to familiarize yourself with their services and resources. Please visit the DAS website for contact information.

If you have already been approved for accommodations, class staff are ready to assist with implementation. Please inform Professor Sarah E Williams (sew@mit.edu) AND Asher Simon (asimon@mit.edu) who will oversee accommodation implementation for this course.

**Academic Integrity**
In this course, I will hold you to the high standard of academic integrity expected of all students at the Institute. I do this for two reasons. First, it is essential to the learning process that you are the one doing the work. I have structured the assignments in this course to enable you to gain a mastery of the course material. Failing to do the work yourself will result in a lesser understanding of the content, and therefore a less meaningful education for you. Second, it is important that there be a level playing field for all students in this course and at the Institute so that the rigor and integrity of the Institute’s educational program are maintained.

Violating the Academic Integrity policy in any way (e.g., plagiarism, unauthorized collaboration, cheating, etc.) will result in official Institute sanction. Possible sanctions include receiving a failing grade on the assignment or exam, being assigned a failing grade in the course, having a formal notation of disciplinary action placed on your MIT record, suspension from the Institute, and expulsion from the Institute for very serious cases.

Please review MIT’s Academic Integrity policy and related resources (e.g., working under pressure; how to paraphrase, summarize, and quote; etc.) and contact me if you have any questions about appropriate citation methods, the degree of collaboration that is permitted, or anything else related to the Academic Integrity of this course.

**Student Support**
**Undergraduate Students: Student Support Services (S3)**
If you are dealing with a personal or medical issue that is impacting your ability to attend class, complete work, or take an exam, you should contact a dean in Student Support Services (S3). S3 is here to help you. The deans will verify your situation, provide you with support, and help you work with your professor or instructor to determine next steps. In most circumstances, you
will not be excused from coursework without verification from a dean. Please visit the S3 website for contact information and more ways that they can provide support.

Website: https://studentlife.mit.edu/s3

Graduate Students: GradSupport

As a graduate student, a variety of issues may impact your academic career including faculty/student relationships, funding, and interpersonal concerns. In the Office of Graduate Education (OGE), GradSupport provides consultation, coaching, and advocacy to graduate students on matters related to academic and life challenges. If you are dealing with an issue that is impacting your ability to attend class, complete work, or take an exam, you may contact GradSupport by email at gradsupport@mit.edu or via phone at (617) 253-4860.

Website: https://oge.mit.edu/development/gradsupport/