Details and Acknowledgements:

Sustainable Development Action Lab
Nepal Study Center (NSC)
University of New Mexico (UNM)

Acknowledgments: UNM's Himalayan Study Abroad Program (2017, 2018 undergraduate student cohorts), Pratiman-Neema Memorial Foundation (local collaborator, Siddharthanagar, Province 5, Nepal), Province 5 Government, Thomas G. Henderson (undergraduate student), and many economics graduate mentors.

1 Undergraduate Honors student, 2 Professor, Department of Economics, UNM & Director of NSC
Thesis Topic and Structure

Topic: Air Quality and Environmental Education in Nepal

- **Literature review**
  - Case studies of areas with similar pollution
  - Practices for communicating data/ citizen science initiatives

- **Analyzing air quality in Province 5 of Nepal**
  - Patterns, predictive modeling

- **Appendices**
Context

- Emphasis: creating tools to communicate with data.
- A mix of analysis and “end solution”
- The sensors are already there, and prior research does indicate that there is a large air quality problem in the district.
- Current focus is on infrastructure for web dashboard, plus exploratory analysis
How to get to a web based dashboard?

What would a useful one even look like?
Starting Point

- A collection of different sensors and types of sensors bringing in data
- An assortment of CSV files floating around, with no standardized way to handle them
Data Pipeline to Dashboard

**Collection**
Retrieving the data from the sensors, once it has been generated.

**Cleaning**
Cleaning and formatting the data so that it can be directly used or added to the database.

**Database**
Where the data is stored. Can be queried, and used for multiple functions.

**Dashboard**
The updating public display of the data. Makes data from the database viewable.
Toolsets

R
  ggplot2
  Analysis

MySQL
  Workbench
  HeidiSQL
  command line

Python
  Pandas
  Bidict
Data Standardization

- What does the data need to look like across sensors?
- Determining time zones, units like temperature, variable names
- Example: Time and Date
  - Consistent time zone, format, field name
  - These were not originally consistent
Sensor Naming Conventions

\(\text{Province, District, City, Ward, Sensor Location, Sensor Type}\)

P5_Parasi_Ramg_NA_Prithvi_PA

Long Version: *Province 5, Nawalparasi, Ramgram, unknown Ward, Prithvi Chandra Hospital, Purple Air*
Cleaning Scripts

Enforce/ Implement the standards defined.

A collection of custom classes and functions

Main modules: name, and format

Automated formatting given sensor type.
Database Structure

MySQL database

Table For Each Sensor

Dictionary Table for Names

Additional as needed

User Permissions

<table>
<thead>
<tr>
<th>Example Table</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time</strong></td>
</tr>
<tr>
<td>Datetime values</td>
</tr>
</tbody>
</table>
Servers and Hosting

- Virtual/Shared Linux Server
- cPanel
  - control panel for accessing the server and editing the site
  - can edit access settings
  - GUI for file management
- UNM Hosting and domain name services
  - Continuity and support
  - Price of scaling is reasonable
- Currently, split between two servers
Site Structure

Dashboard:
A collection of graphs and interpretations that update based on the data in the database. Could be made more interactive in the future.

Analysis:
Static graphs and interpretation, informational pages, papers. An example would be looking at historical trends, or uploading a research paper.
Dashboard and Analysis

https://dempnsc.unm.edu
Developing a codebase for standardized graphs

General analysis work
Technical Future

My work:
- Automation
- User interface/useability
- Error handling
- Some feature expansion

Possible extensions:
- Dashboard expansion
- Geospatial
- Mobile Application
- Traditional Analysis
Questions?