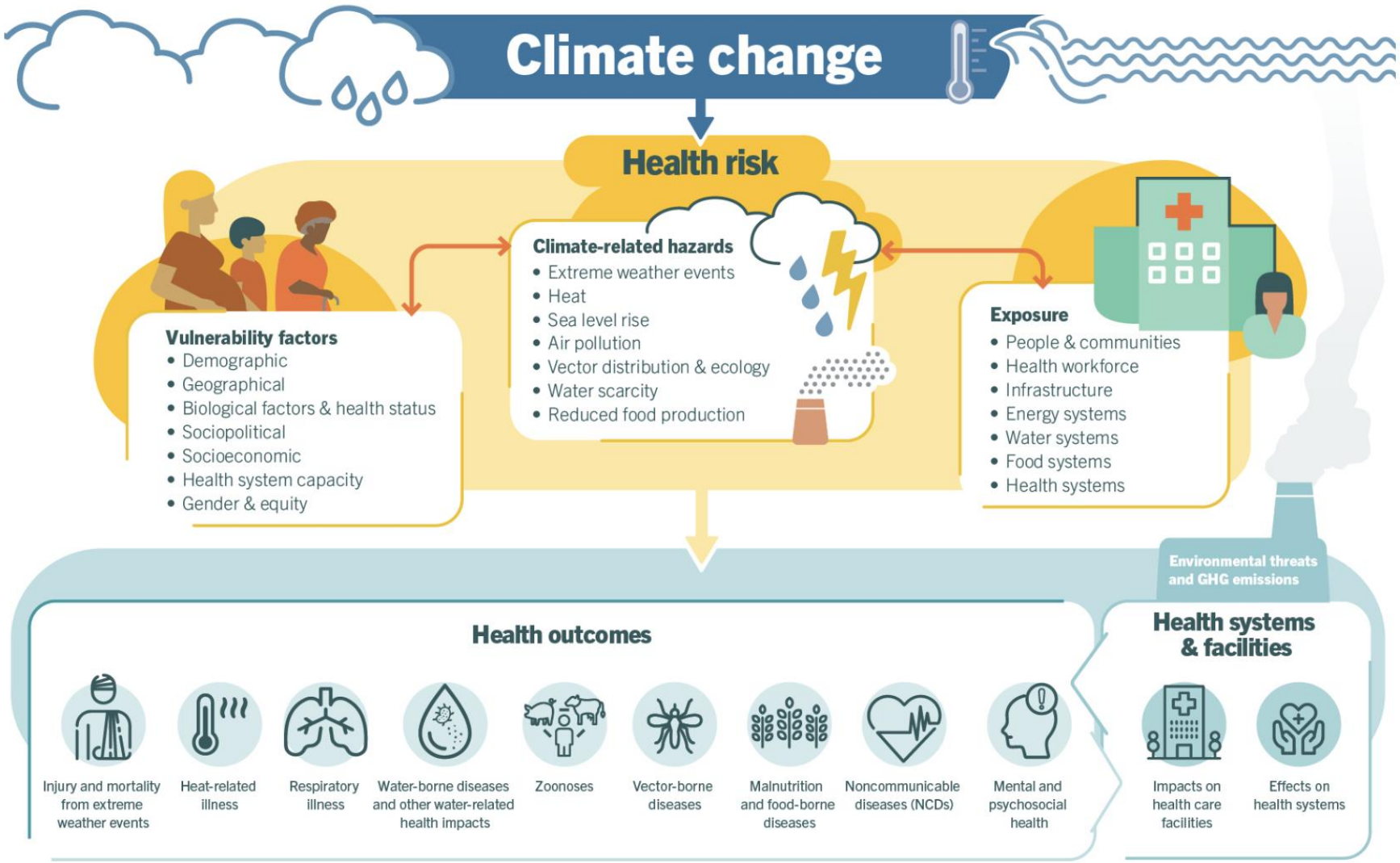
Effects of Desertification on Health-Related Outcomes

Group 2: Andrea Stofko, Courtney Gauchel, Liam Wilson, Maya Wheeler, and Spencer Gilbert

Project Overview

- Background and Motivation
- Project Objectives
- Datasets
- Data Processing
- Data Integration
- Limitations



(Figure found on WHO climate change factsheet https://www.who.int/news-room/fact-sheets/detail/climate-change-and-health)

Background and Motivation

- **Desertification:** The process that occurs when previously fertile land degrades into desert.

- The World Health Organization (WHO) has identified multiple sources of desertification that have caused this process to accelerate throughout the 20th and 21st century.
    - Agricultural and livestock overproduction, urbanization, deforestation, and extreme weather events.

- Climate change has led to rising temperatures and decreases in annual precipitation, resulting in the desertification of regions worldwide.

- Desertification can have massive impacts on human health. According to the WHO, areas that have been affected by desertification have higher threats of:
    - Malnutrition, more water- and food-borne diseases, **an increase in respiratory diseases**, and an increase in the spread of infectious diseases.

Project Objectives

- Our project analyzes trends in mortality from respiratory diseases and heat-related illnesses that are linked to rising temperatures, prolonged drought, and increased wildfire frequency from 2000 to 2020.
- We chose to create a database focused on trends in the coastal western US due to the notable changes in annual temperature, wildfire prevalence, and the region's high population density.
- **Primary research questions:**
  - How has the frequency and severity of extreme heat events, droughts, and wildfires changed in the coastal western US during this time?
  - How have these climate factors changed mortality rates associated with asthma, COPD, and heat stroke?

Data Granularity

- Temporal:
  - Monthly data from 2000 to 2020
- Geospatial:
  - Counties on the West Coast (California, Oregon, and Washington)

Datasets- Wildfires

- Data needed: county, year/month, wildfire acres
- California:
  - California Fire Perimeter Dataset (California Department of Forestry and Fire)
    - Date and geospatial shape (GeoJson) data per fire
- Oregon:
  - ODF Fire Occurrence Dataset (Oregon Department of Forestry)
    - Date, acreage, and county: trivial reconciliation
- Washington
  - Washington Large Fires Dataset (Washington State Department of Natural Resources)
    - Date and geospatial shape data (GeoJson) per fire

Datasets- Health Outcomes

- Data needed: county, year/month, deaths
- Asthma, COPD, and hyperthermia mortality data:
  - CDC Wonder Database
    - Query results for California, Oregon, and Washington counties using ICD codes for asthma (J45), COPD (J44), and hyperthermia (X30)

Datasets- Drought/Temperature

- Data needed: county, year/month, max temperature, and total precipitation
- Monthly precipitation and max temperature
  - OSU PRISM Data
    - 4km gridded max temperature and precipitation per month

Data Processing: Health

- Data was obtained as flat files (CSVs) from the publicly available state databases
- Datasets were created using consistent temporal and geospatial constraints, with monthly rates from 2000 to 2020 at the county level
- Python was used to process the CSVs and create consistent formatting
- Data types were converted as needed

Data Processing: Climate

- Wildfire data
  - CSVs (Oregon ODF)
  - GeoJSON (Washington DNR; CalFire)
  - County shapefiles
- Python (geopandas; pygris) used for geospatial intersection calculation and obtention of geospatial county shapes
- Python (pandas) was used to process non-geospatial attributes and create consistent formatting (columns ~ SQL attributes)
- Temp/drought: Oregon PRISM yearly raster zip files (FTP via Finder)
  - Subdirectories for monthly PRISM totals for precipitation and max temperature
  - Rasterio, Geopandas, and pygris libraries were used to mask country raster files to counties
  - Max raster value was taken for temperature per month for county temperature maximums
  - Summed value was taken for precipitation totals (totaling rainfall per county)

Data Processing Challenges

Data inconsistencies:

- Reconciled geographic and temporal inconsistencies across multiple wildfire datasets to ensure that formatting was standardized for downstream analysis
- Example: The Oregon wildfire dataset was mapped using raw coordinate data. This was converted to the county-level format using geospatial mapping tools (GeoPandas library).

Missingness:

- For datasets that contained missing values, these rows were identified and dropped.
- Example: Some wildfire datasets included rows with missing months. These were removed.
- Heatmaps were used to visualize any remaining missingness and evaluate data completeness before analysis.

Data Integration

A **relational database** was constructed to allow for efficient data retrieval and cross-referencing of environmental conditions with health outcomes to enable a deeper analysis of the effects of desertification on public health.

SQL tables were created for the database to compile data from each of the cleaned datasets:

- PRISM table– includes monthly drought severity indices, average temperatures, heat wave events by county
- Wildfires table – stores total monthly county-level wildfire acreage
- Mortality table – contains mortality counts for heat-related illness, asthma, and COPD by month and county

These tables can be queried and joined using the county codes and names.

Limitations

- CDC mortality data will appear as "suppressed" if there are fewer than 10 deaths per month in a given county
- This is an issue more in the rural counties, where # of deaths are a lot smaller than urban counties
- Potential Solution: Pool data from multiple ICD codes to get less "suppressed" values

California	Shasta County, CA	6089	2019.0	8.0	Suppressed
California	Shasta County, CA	6089	2019.0	9.0	16
California	Shasta County, CA	6089	2019.0	10.0	16
California	Shasta County, CA	6089	2019.0	11.0	Suppressed
California	Shasta County, CA	6089	2019.0	12.0	17
California	Shasta County, CA	6089	2020.0	1.0	14

California	Amador County, CA	6005	2017.0	4.0	0
California	Amador County, CA	6005	2017.0	5.0	0
California	Amador County, CA	6005	2017.0	6.0	Suppressed
California	Amador County, CA	6005	2017.0	7.0	0
California	Amador County, CA	6005	2017.0	8.0	0

Questions?