Berkeley Earth + Emission EDA

Team 5
Pascual Eley, Victor Yin, Erik Sambrailo

Berkeley
UNIVERSITY OF CALIFORNIA

Datasets

Berkeley Earth

Time series regional average temperature datasets
237 countries with potentially 192 years of data
https://berkeleyearth.org/data/

UN Emissions Data

Greenhouse gas emissions datasets for 7 types
of emissions for 43 countries over 30 years
http://data.un.org/Data.aspx?d=GHG&f=seriesID%3aCH4#GHG

Questions

- How have temperatures changed globally over the past couple centuries?
  - How does this differ by country?
  - Does proximity to the equator have a correlation on temperature change?

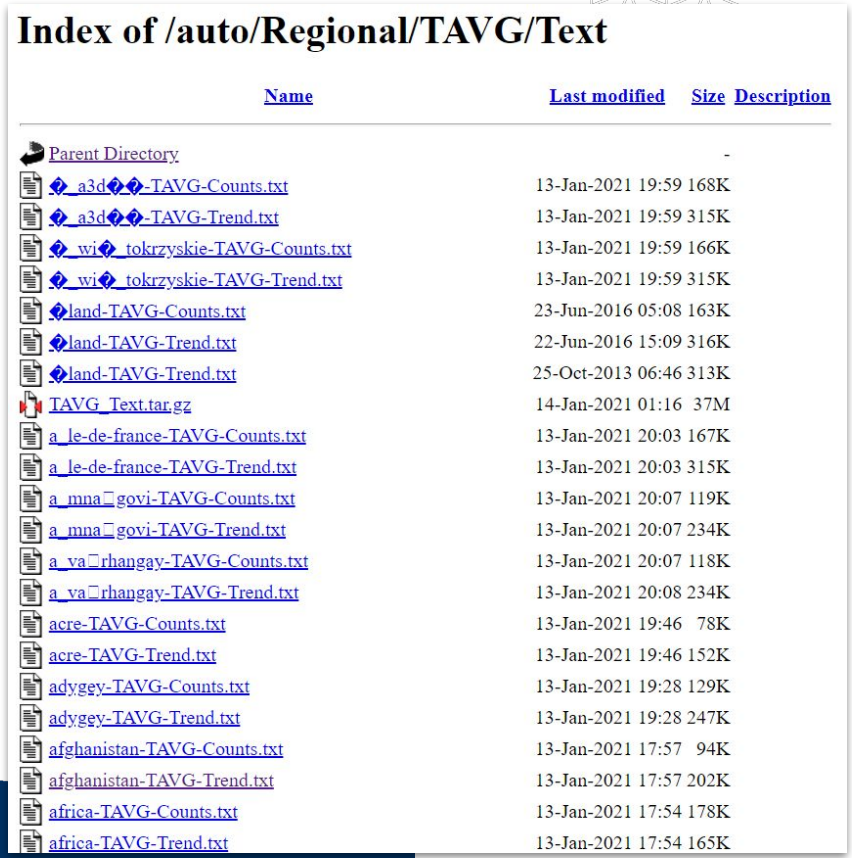
- How have greenhouse gas emissions changed over time?
  - How does this differ by country?
  - How does it differ by type of gas?
  - Does a country's emission have any correlation to its change in temperature?

Scraping Berkeley Earth

Downloading Country data

- Iterated through list of countries
- download .txt files from directory that matches

Only (4) countries of 237 that were not downloadable.



Index of /auto/Regional/TAVG/Text

Name	Last modified	Size	Description
Parent Directory		-	
[ICO]	13-Jan-2021 19:59	168K	
[ICO]	13-Jan-2021 19:59	315K	
[ICO]	13-Jan-2021 19:59	166K	
[ICO]	13-Jan-2021 19:59	315K	
[ICO]	23-Jun-2016 05:08	163K	
[ICO]	22-Jun-2016 15:09	316K	
[ICO]	25-Oct-2013 06:46	313K	
[ICO]	14-Jan-2021 01:16	37M	
[ICO]	13-Jan-2021 20:03	167K	
[ICO]	13-Jan-2021 20:03	315K	
[ICO]	13-Jan-2021 20:07	119K	
[ICO]	13-Jan-2021 20:07	234K	
[ICO]	13-Jan-2021 20:07	118K	
[ICO]	13-Jan-2021 20:08	234K	
[ICO]	13-Jan-2021 19:46	78K	
[ICO]	13-Jan-2021 19:46	152K	
[ICO]	13-Jan-2021 19:28	129K	
[ICO]	13-Jan-2021 19:28	247K	
[ICO]	13-Jan-2021 17:57	94K	
[ICO]	13-Jan-2021 17:57	202K	
[ICO]	13-Jan-2021 17:54	178K	
[ICO]	13-Jan-2021 17:54	165K	

Berkeley
UNIVERSITY OF CALIFORNIA

Converting .txt files to DF

- Consistent formatting of .txt files
- Iterated through files to compile dataframe

```
% that month (rounding down if the center is in between months).  For e
% the annual average from January to December 1950 is reported at June
%
% Values are reported as missing (i.e. NaN) when station coverage withi
% the region becomes too low, even though a limited number of observati
% still have been made.  Time averages over intervals with some missing
% be reported as long as at least 75% of the necessary values are avail
%
%                    Monthly          Annual          Five-year          Te
% Year, Month,  Anomaly, Unc.,  Anomaly, Unc.,  Anomaly, Unc.,  Anom

  1848     5    -0.297  2.037       NaN    NaN       NaN    NaN        N
  1848     6    -0.796  2.136       NaN    NaN       NaN    NaN        N
  1848     7    -0.113  1.937    -0.777  0.639       NaN    NaN        N
  1848     8    -0.462  1.937    -0.743  0.644       NaN    NaN        N
  1848     9    -1.272  1.865    -0.676  0.669       NaN    NaN        N
  1848    10    -0.934  1.880    -0.712  0.687       NaN    NaN        N
  1848    11    -0.769  1.835    -0.723  0.663       NaN    NaN        N
  1848    12    -1.842  2.034    -0.742  0.663       NaN    NaN        N
  1849     1    -0.095  2.550    -0.731  0.685       NaN    NaN        N
  1849     2    -0.560  2.098    -0.704  0.711       NaN    NaN        N
  1849     3    -0.367  2.068    -0.718  0.714       NaN    NaN        N
  1849     4    -1.041  2.307    -0.688  0.713       NaN    NaN        N
  1849     5    -0.427  1.894    -0.706  0.739       NaN    NaN        N
  1849     6    -1.016  1.976    -0.699  0.676       NaN    NaN        N
  1849     7     0.012  2.050    -0.855  0.641       NaN    NaN        N
  1849     8    -0.141  1.952    -0.837  0.650       NaN    NaN        N
  1849     9    -1.431  1.843    -0.811  0.661       NaN    NaN        N
  1849    10    -0.584  1.845    -0.818  0.718       NaN    NaN        N
```

Berkeley
UNIVERSITY OF CALIFORNIA

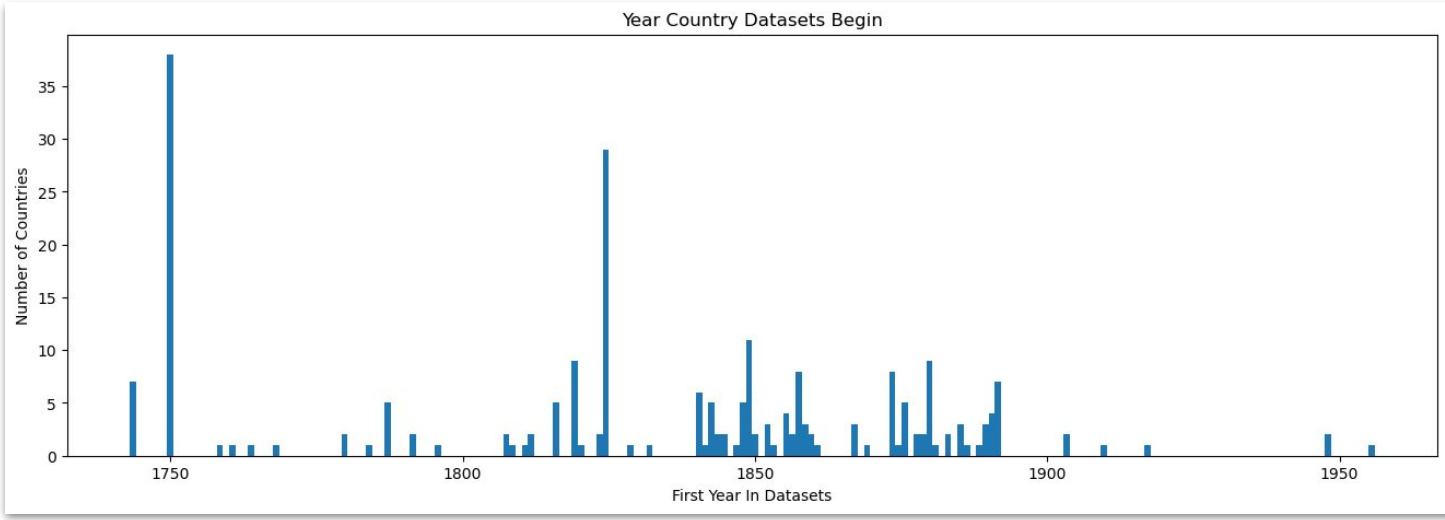
Dataset Overview & Assumptions

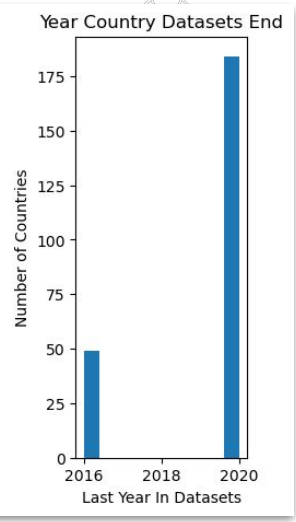
- Country temperature baseline
  - Average temp. from Jan 1951–Dec 1980
  - **Assumption:** These values are accurate representations

- Average temperature anomalies, w/ measured uncertainties
  - centered moving averages for time periods of:
    month, & 1–, 5–,10–, 20–Year
  - **Assumption:** Assuming anomalies are accurate and not reviewing uncertainty values

Temp Date Availability

Reviewing start and end dates* for all countries



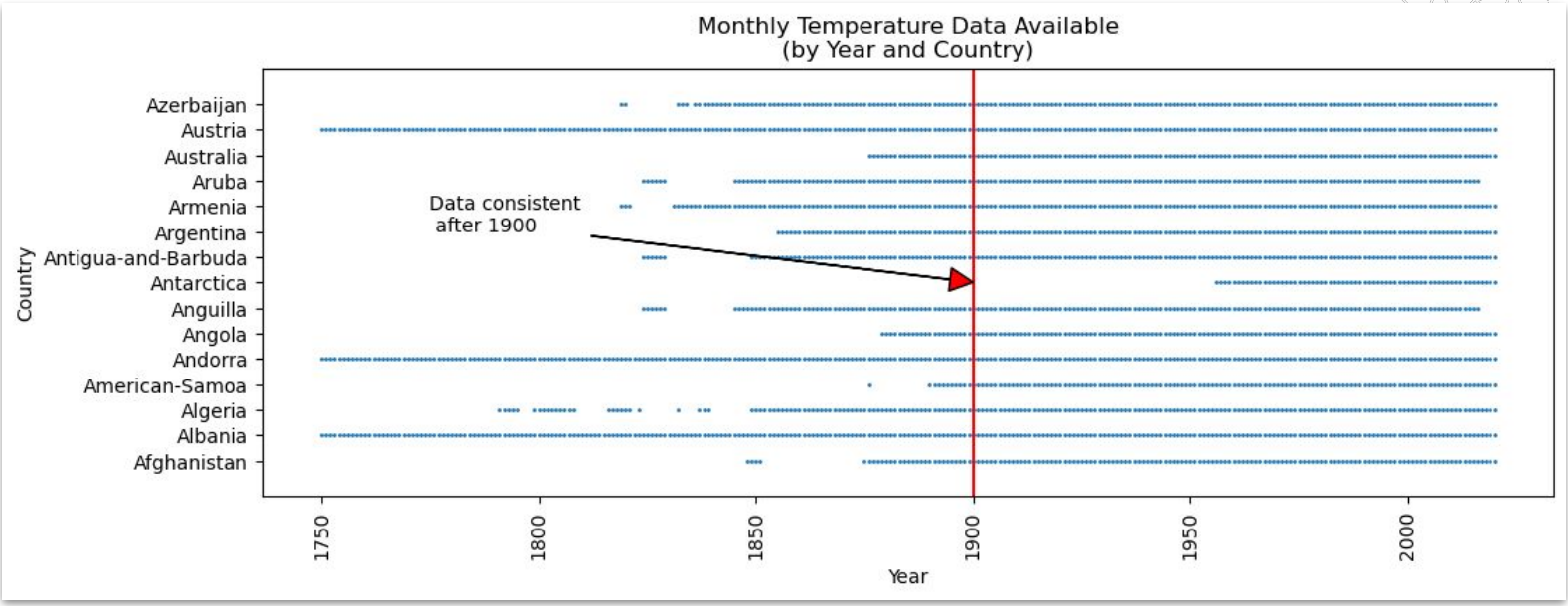


*reviewing monthly temperature anomaly data.

Berkeley
UNIVERSITY OF CALIFORNIA

Temp Date Availability

Reviewing gaps in temperature data



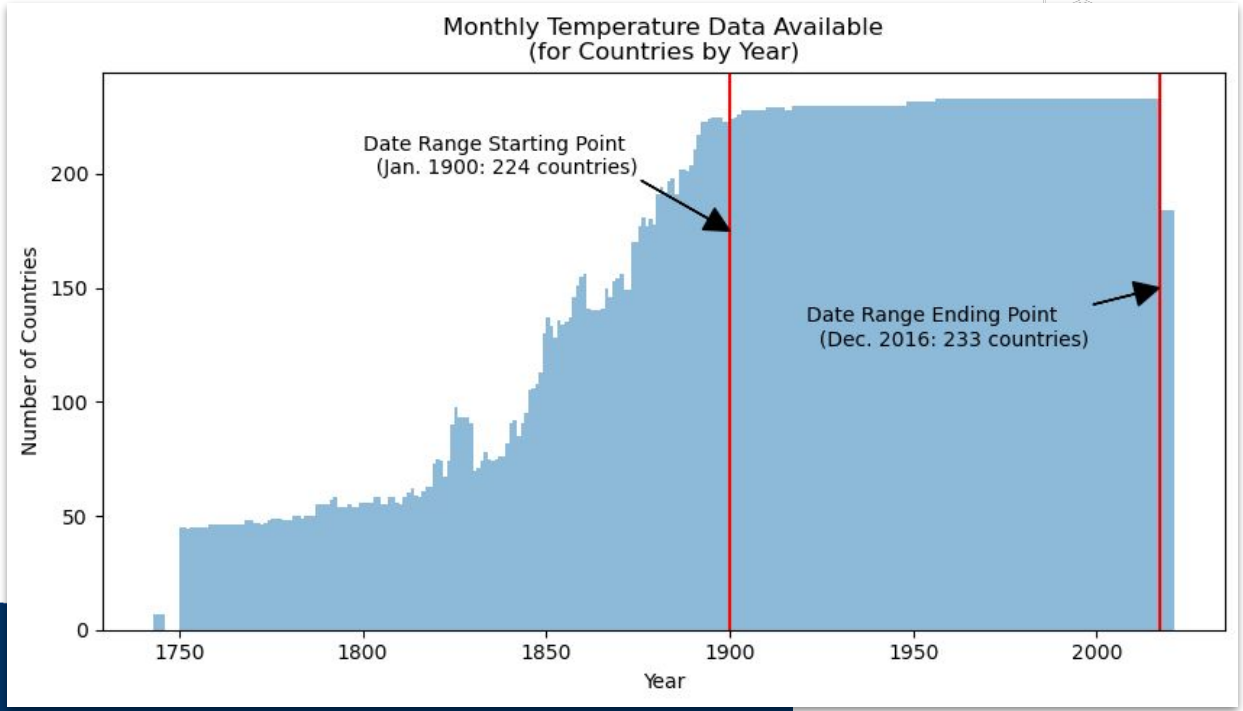
Berkeley
UNIVERSITY OF CALIFORNIA

Temp Date Availability

Identifying number of countries available by year

Countries Added
(between 1900 & 2016)

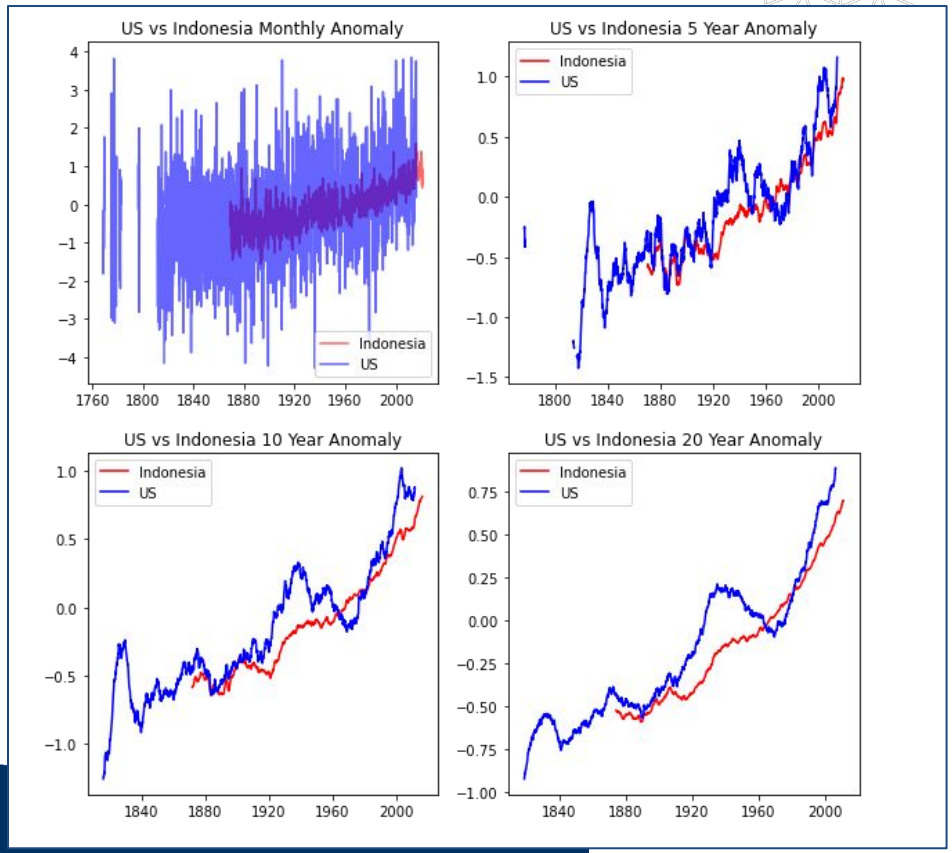
Antarctica
Federated-States-of-Micronesia
French-Southern-and-Antarctic-Lands
Guam
Heard-Island-and-McDonald-Islands
Northern-Mariana-Islands
Palau
Papua-New-Guinea
Solomon-Islands



Berkeley
UNIVERSITY OF CALIFORNIA

Various Anomaly Timeframes

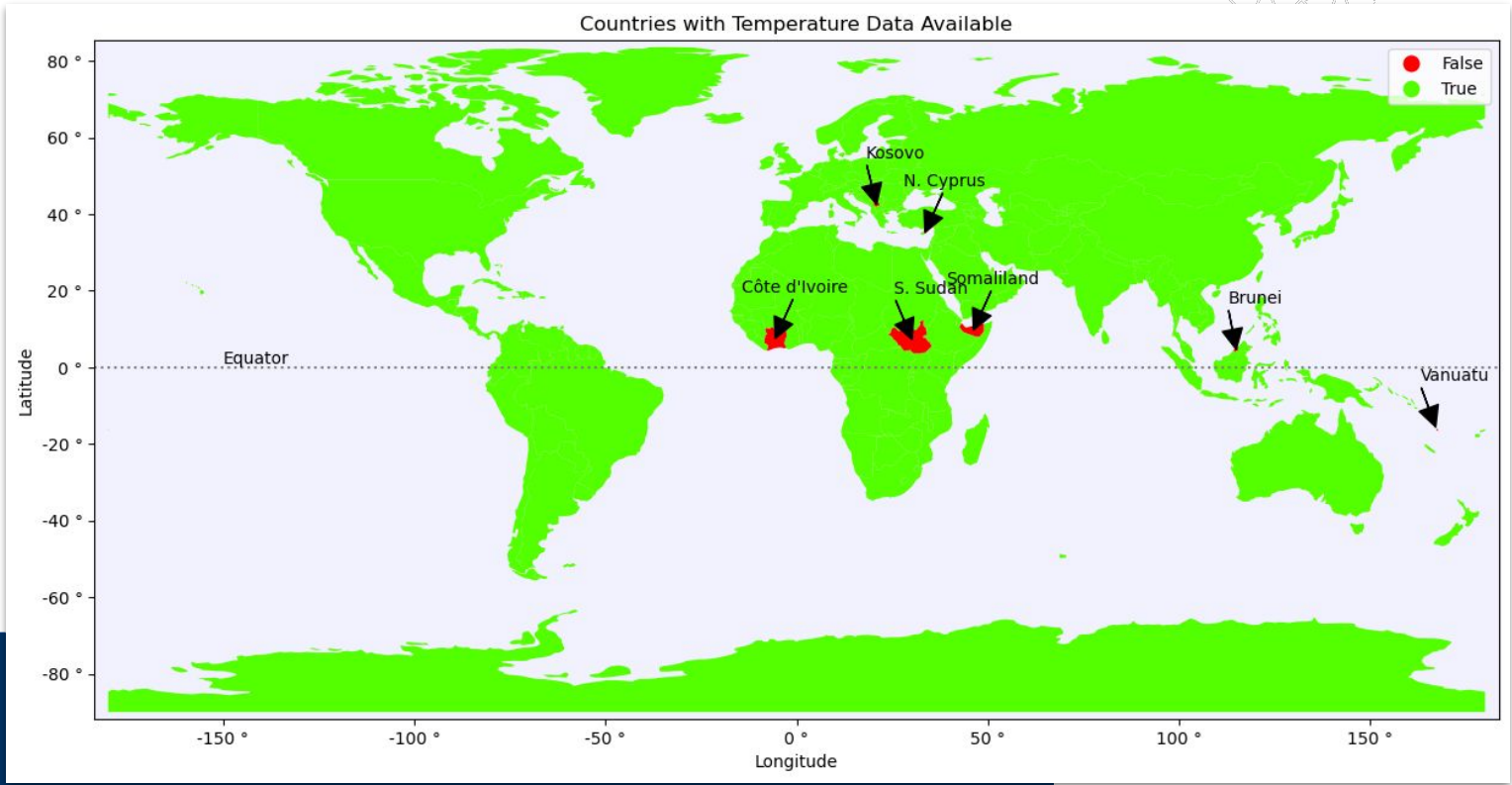
- Used US vs Indonesia as example
- Monthly anomaly – High variability
- 20 Year – Least variability



Berkeley
UNIVERSITY OF CALIFORNIA

Geospatial Data

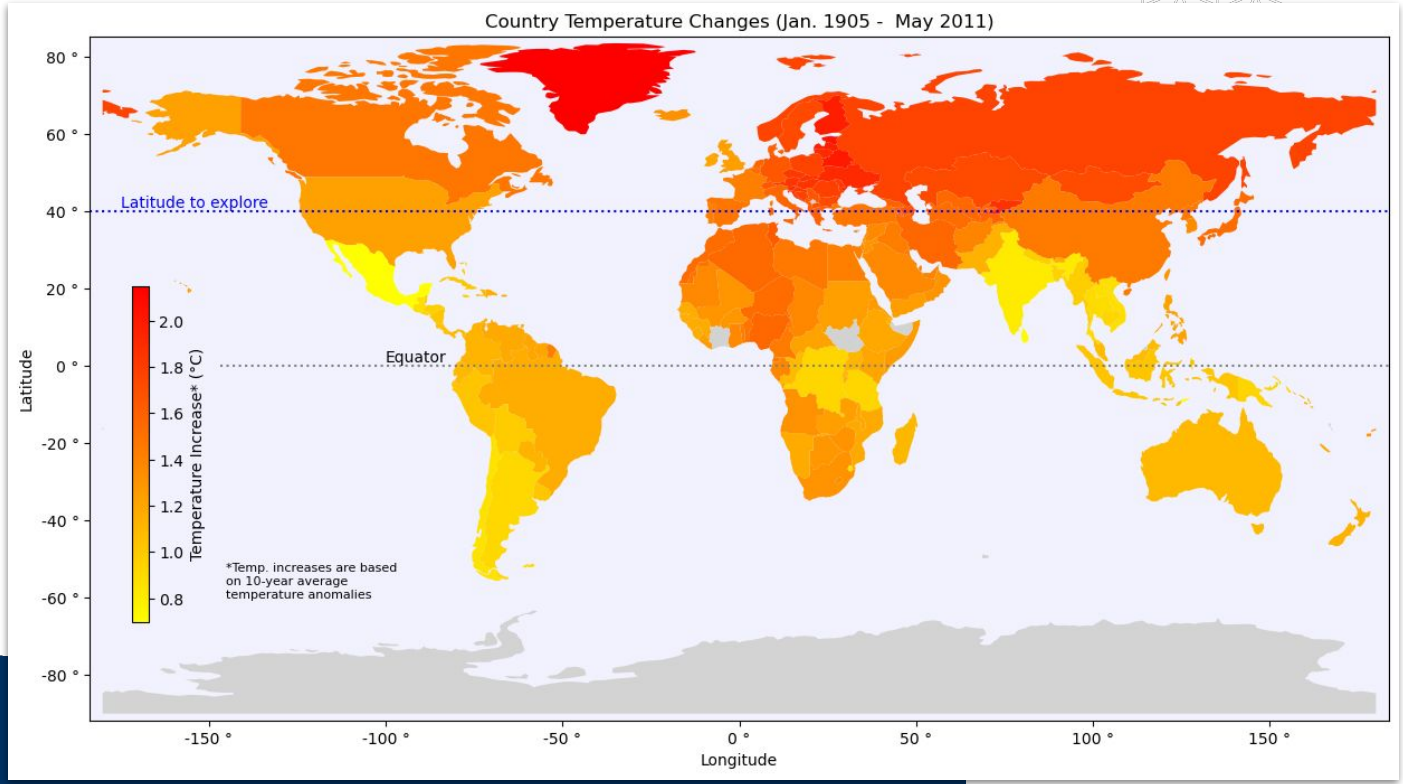
Countries available for geospatial mapping



Berkeley
UNIVERSITY OF CALIFORNIA

Geospatial Temperature Changes

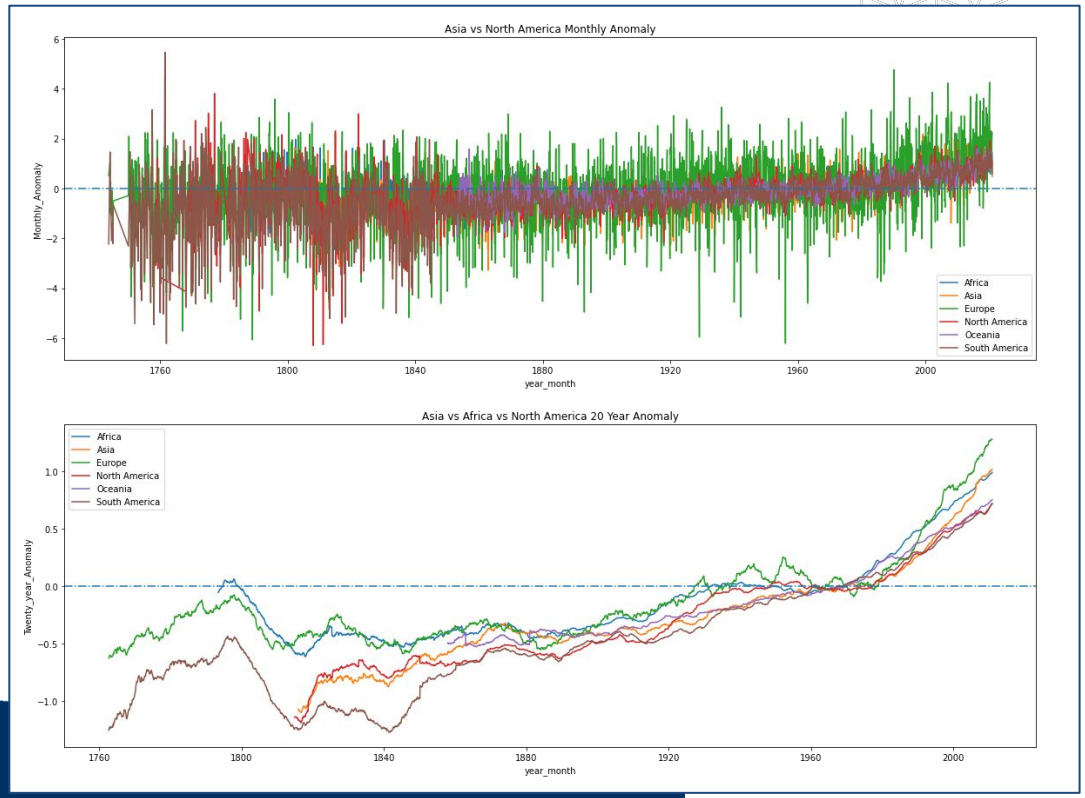
Reviewing geospatial patterns in temperature change



Berkeley
UNIVERSITY OF CALIFORNIA

Anomalies by Region

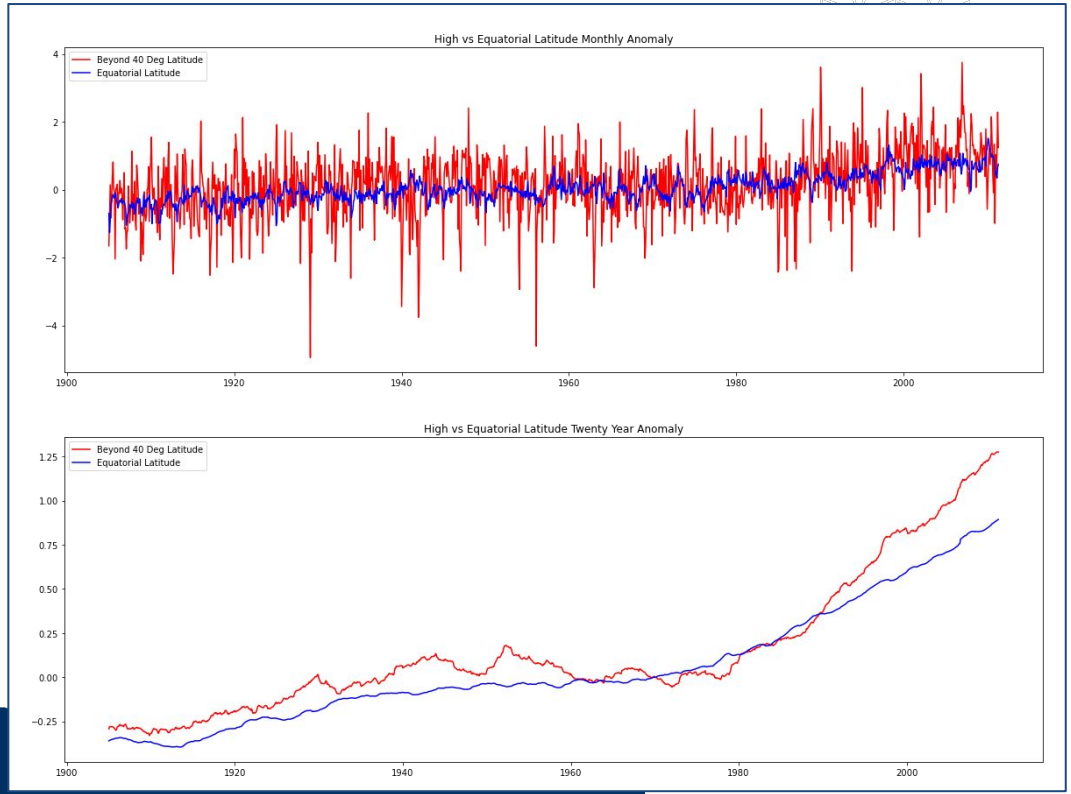
- Compared regions as broken down in our dataset
- Monthly has high variability, but Europe region noticeably higher
- 20 Year trend looks similar across regions



Berkeley
UNIVERSITY OF CALIFORNIA

Latitude Correlation

- Used 40 degrees since it looked to split the hemisphere on the globe
- Monthly anomaly for high latitudes is more variable
- 20 year still variable fo high latitudes, and higher anomaly overall



Berkeley
UNIVERSITY OF CALIFORNIA

Geospatial Temperature Conclusion

Various prior studies have shown that temperature change is greater at the Earth's poles.[1]

The reasons predicted for this are varied:

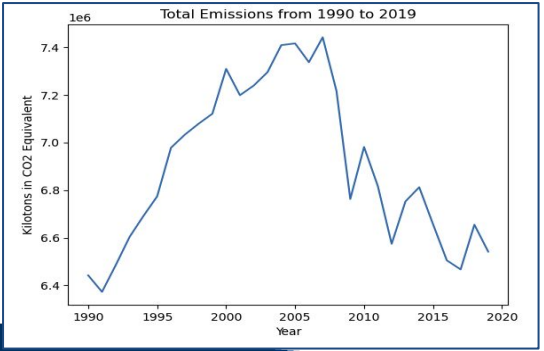
- caused by melting ice which is reducing reflected energy.[1]
- energy in the atmosphere transported to the poles through large weather systems.[1]

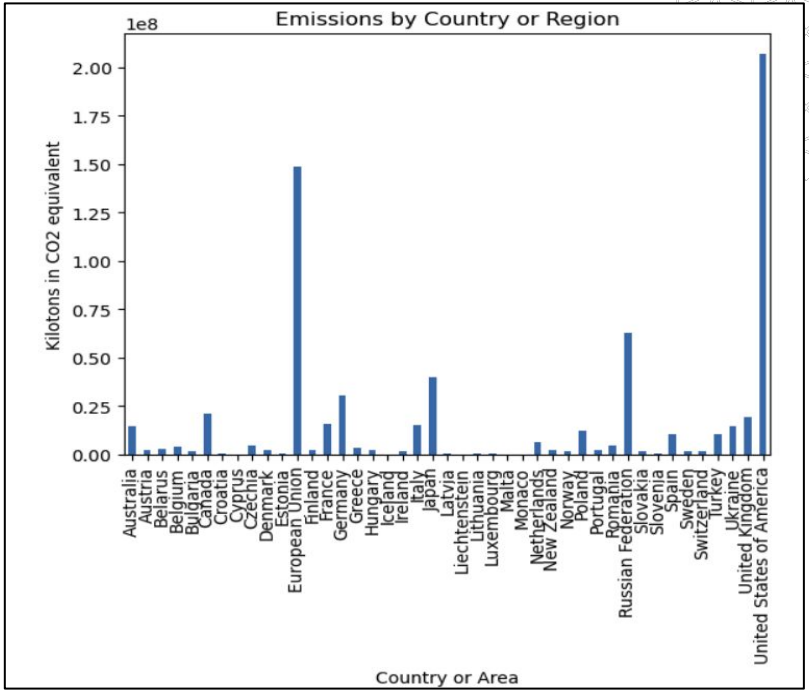
[1] https://www.nasa.gov/topics/earth/features/warmingpoles.html

Emission Data

Emissions by Country

Type	Country or Area	
CO2	United States of America	167927361.97
	European Union	120823499.60
	Russian Federation	49909669.59
	Japan	36938568.58
	Germany	26251704.63
methane	United States of America	21128317.05
	European Union	16550898.53
CO2	Canada	16332753.49
	United Kingdom	15751101.26
N2O	United States of America	13703510.56
CO2	Italy	13001643.35

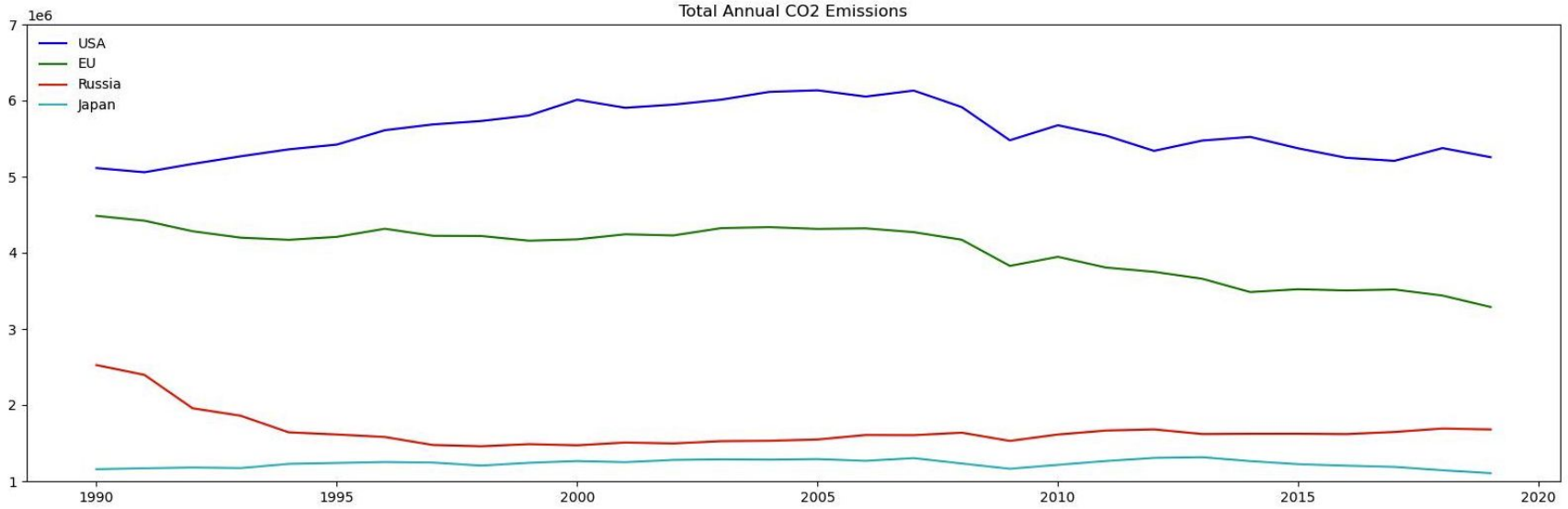




Berkeley
UNIVERSITY OF CALIFORNIA

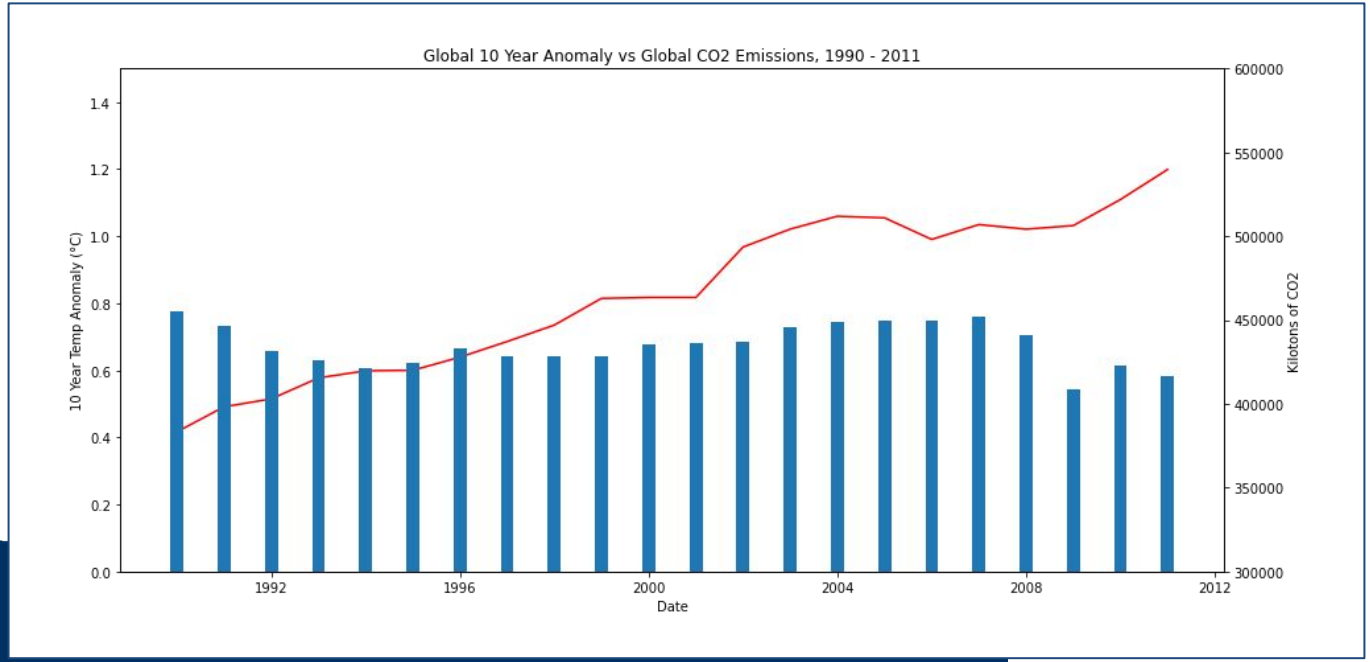
Emission Data

Largest CO2 emitters



Berkeley
UNIVERSITY OF CALIFORNIA

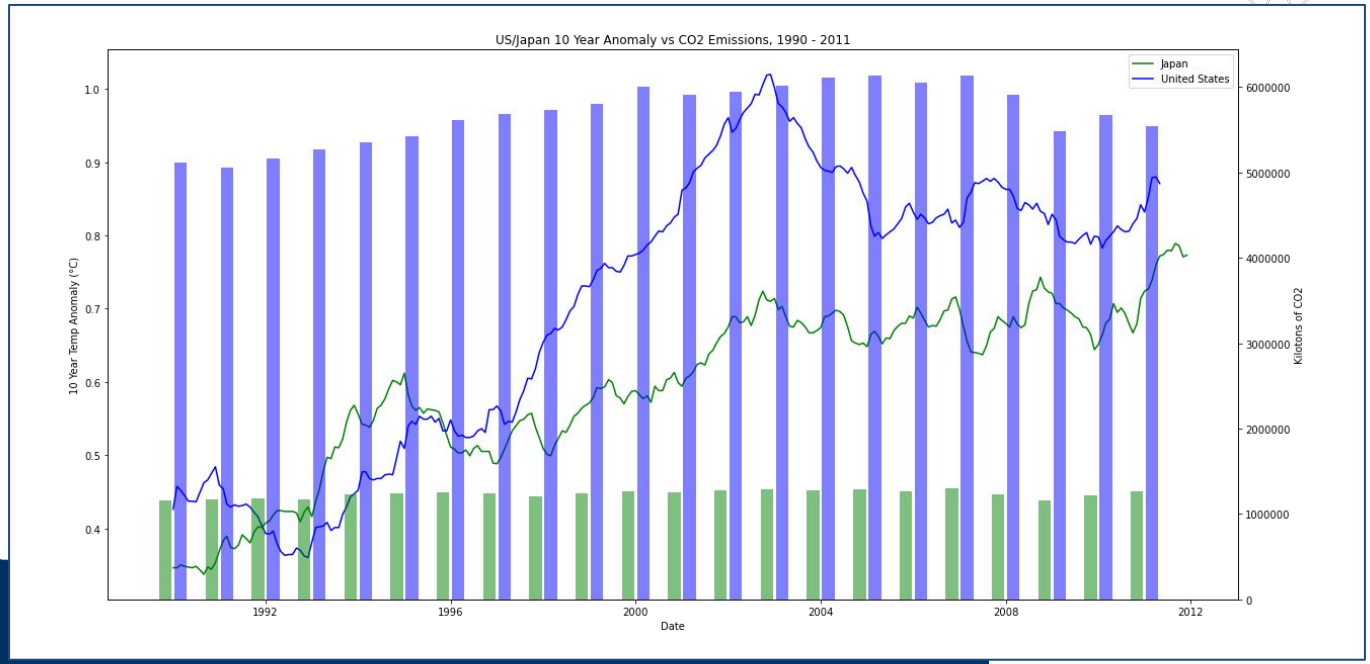
Comparing Emissions to Temperature - Global Average



Berkeley
UNIVERSITY OF CALIFORNIA

Comparing Emissions to Temperature - Japan and US

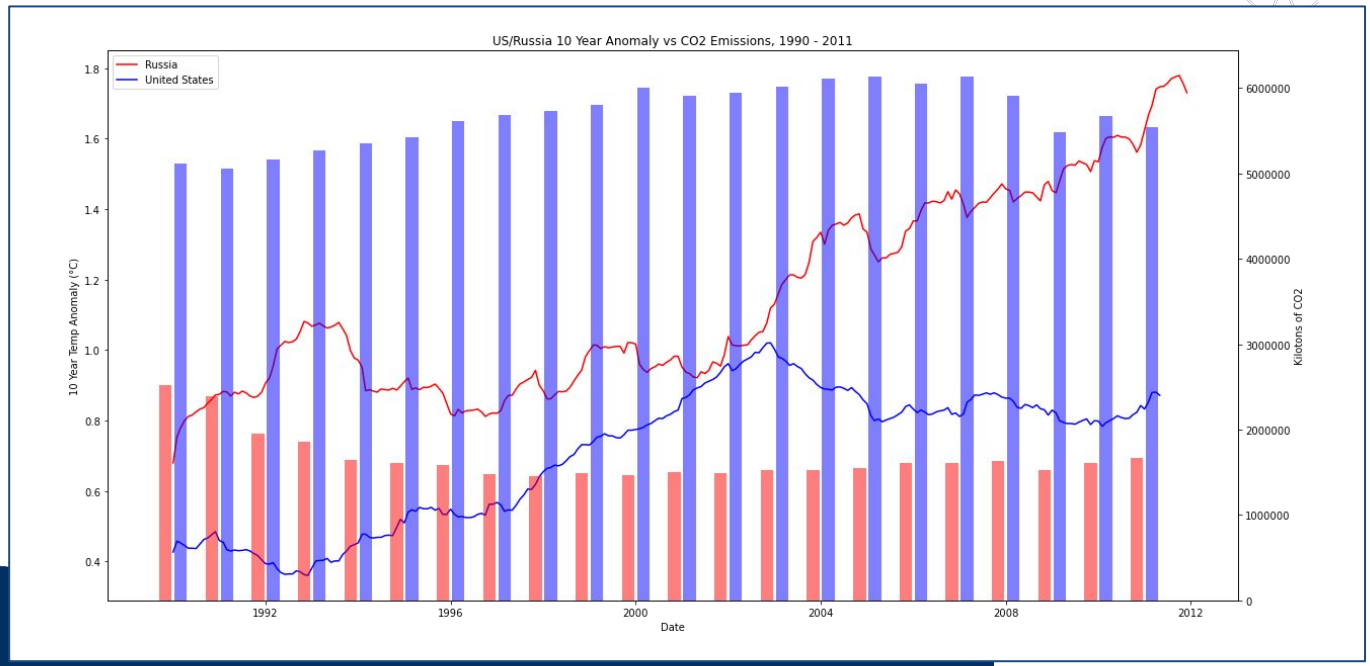
Largest CO2 emitters



Berkeley
UNIVERSITY OF CALIFORNIA

Comparing Emissions to Temperature - Russia and US

Largest CO2 emitters



Berkeley
UNIVERSITY OF CALIFORNIA

Thank you!

Berkeley
UNIVERSITY OF CALIFORNIA