The Statistics of Climate Change

Katie Zhou, Morgan Pruchniewski, Scott Burstein

Introduction

RESEARCH QUESTION: Is there significant evidence to support the existence of climate change?

1. Is there evidence to suggest a statistically significant increase in mean earth surface temperature from early-20th-century levels to what the data show for more recent years?
2. Is the earth changing/increasing temperature at a faster rate now than it was in the early 20th century?
3. Do the data provide evidence of a greater degree of net (positive or negative) year-to-year fluctuation after 1980 than before 1980?
4. Has North America experienced a greater change in annual average temperatures from the first half of the 20th century to 2010 than other continents? Has Europe?

Introduction Continued

THE DATASETS: The data was originally compiled by the Berkeley Earth Data Lab from 16 pre-existing archives, and it was updated to Kaggle in 2017.

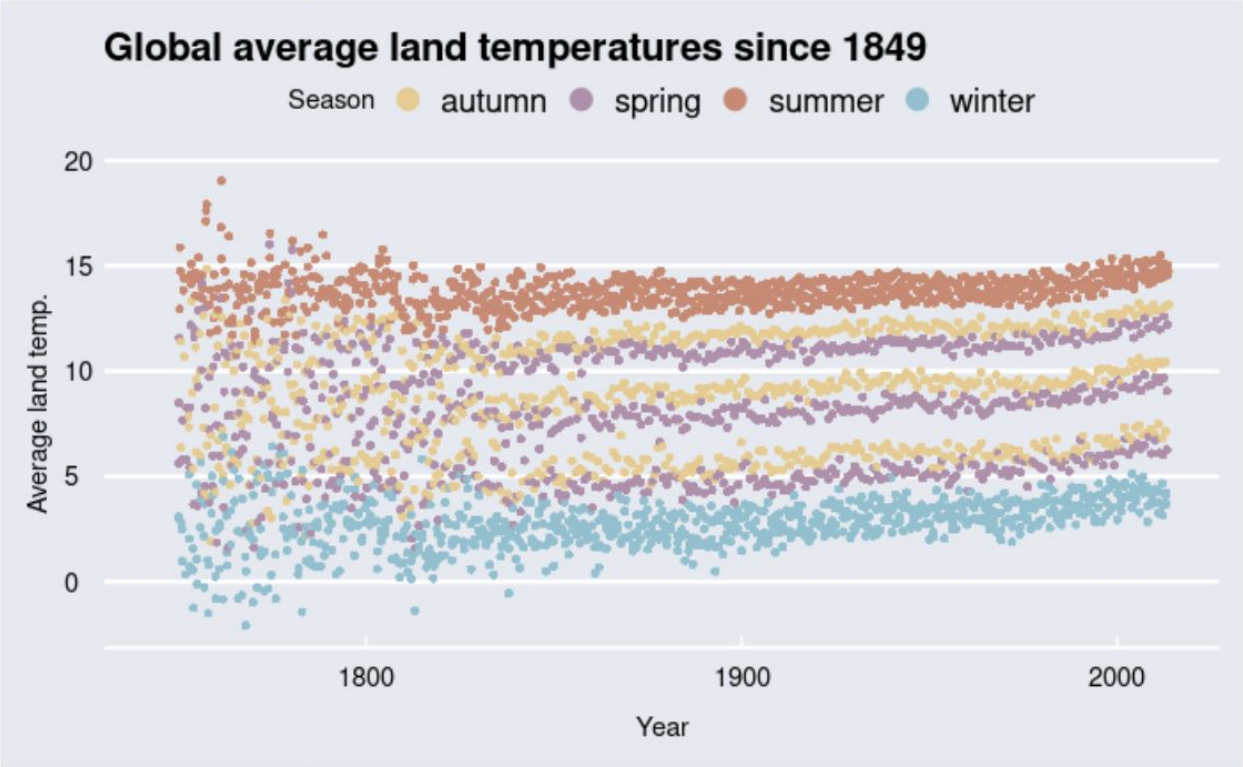
GlobalLandTemperaturesbyMajorCity

- ➔ 239,177 cases of cities and information about them
- ➔ Relevant variables included:
  - ◆ dt (date and time, a discrete numeric)
  - ◆ AverageTemperature (a continuous numeric)
  - ◆ city (nominal categorical
  - ◆ latitude (continuous numeric)
  - ◆ longitude (continuous numeric)

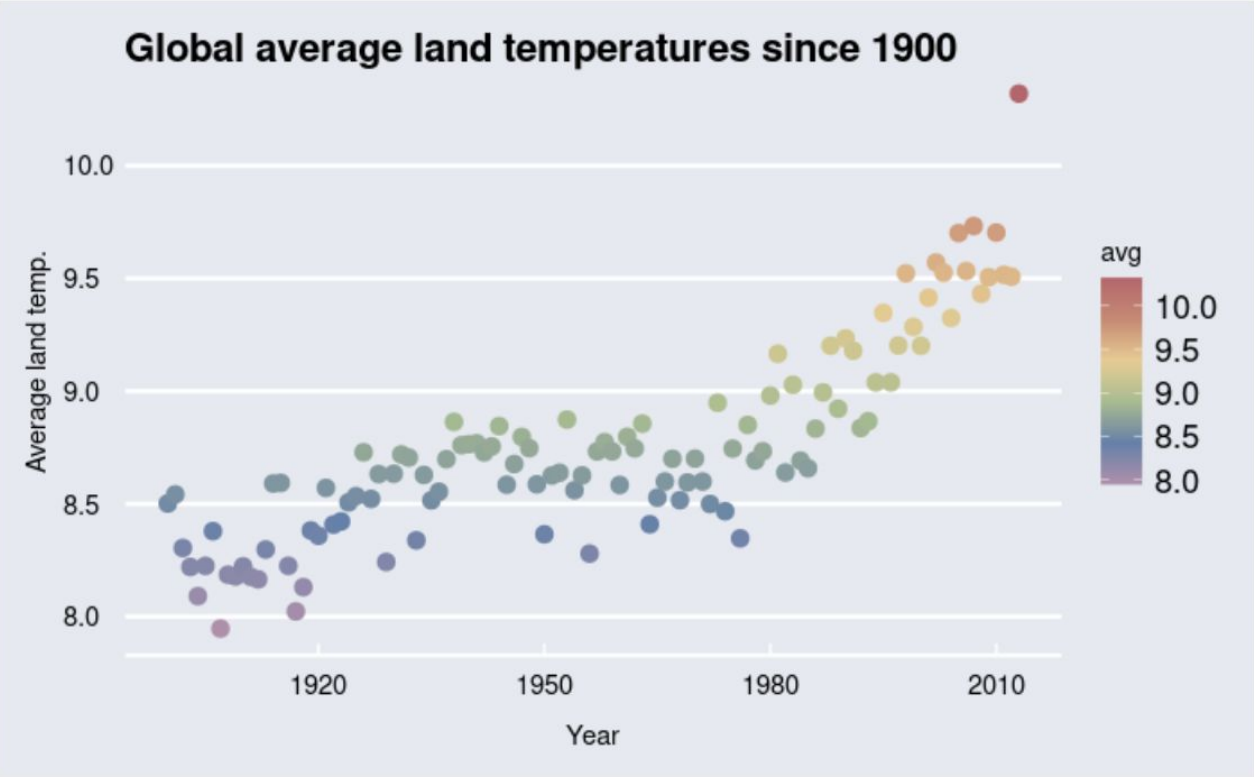
GlobalTemperatures

- ➔ 3,192 cases of recorded global temperature values
- ➔ Relevant variables included:
  - ◆ LandAverageTemperature (continuous numeric)
  - ◆ LandMaxTemperature (a continuous numeric)
  - ◆ LandAndOceanAverageTemperature (a continuous numeric)

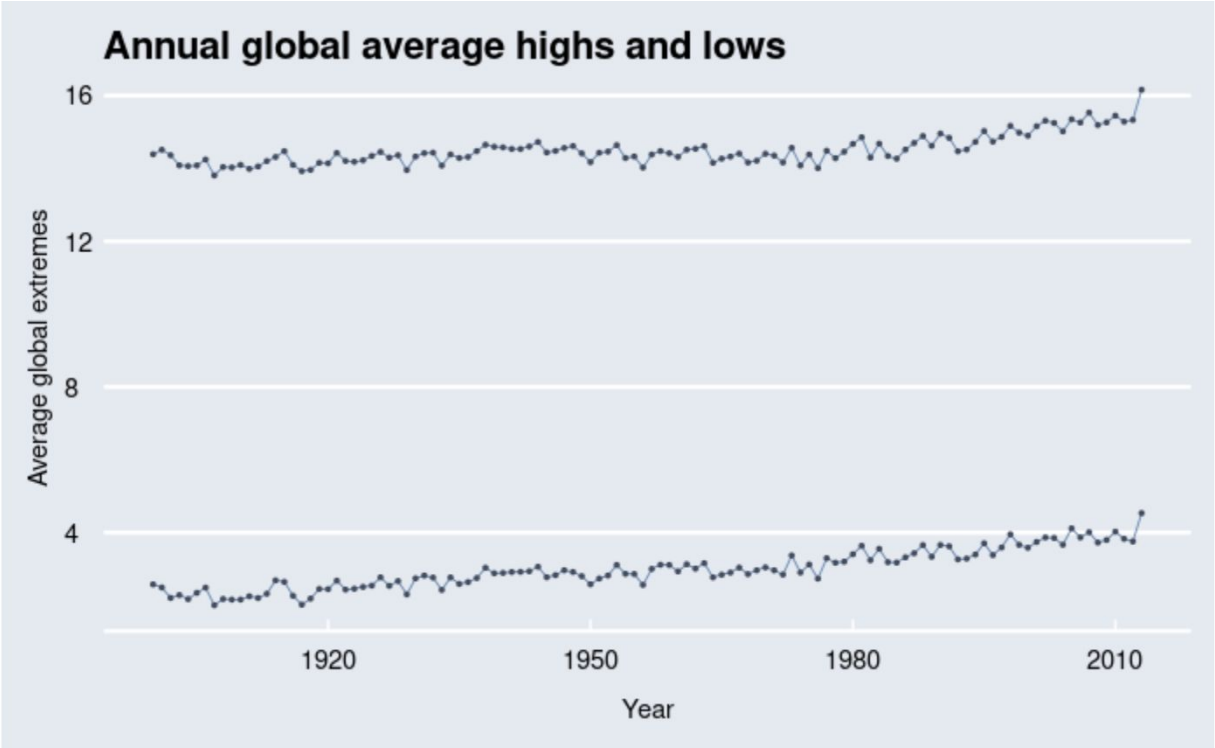
Visualizations



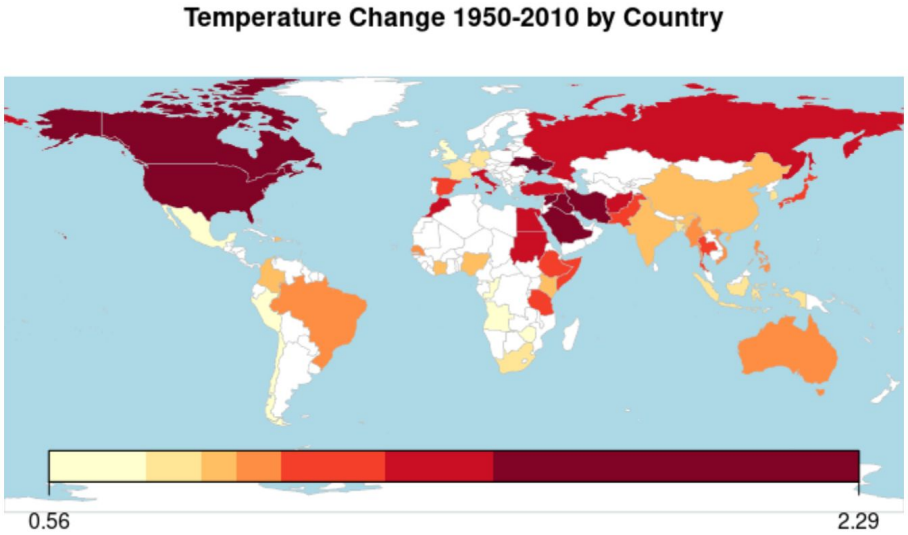
Visualizations

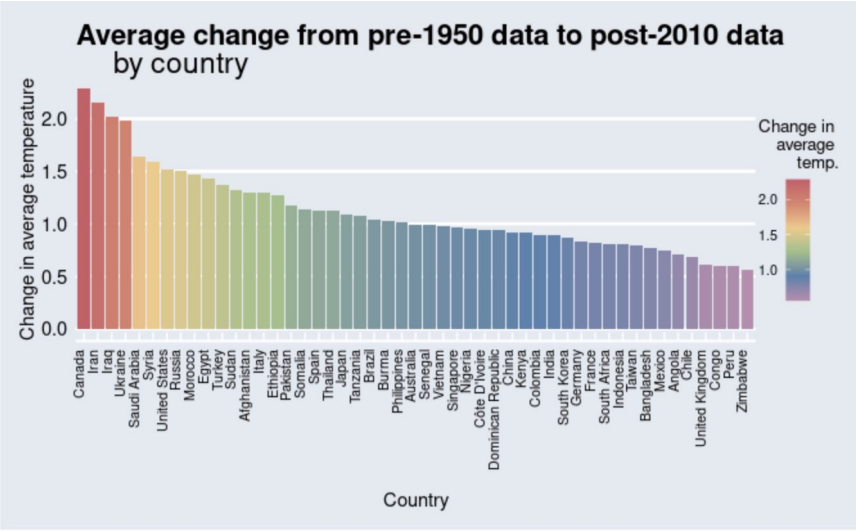


Visualizations

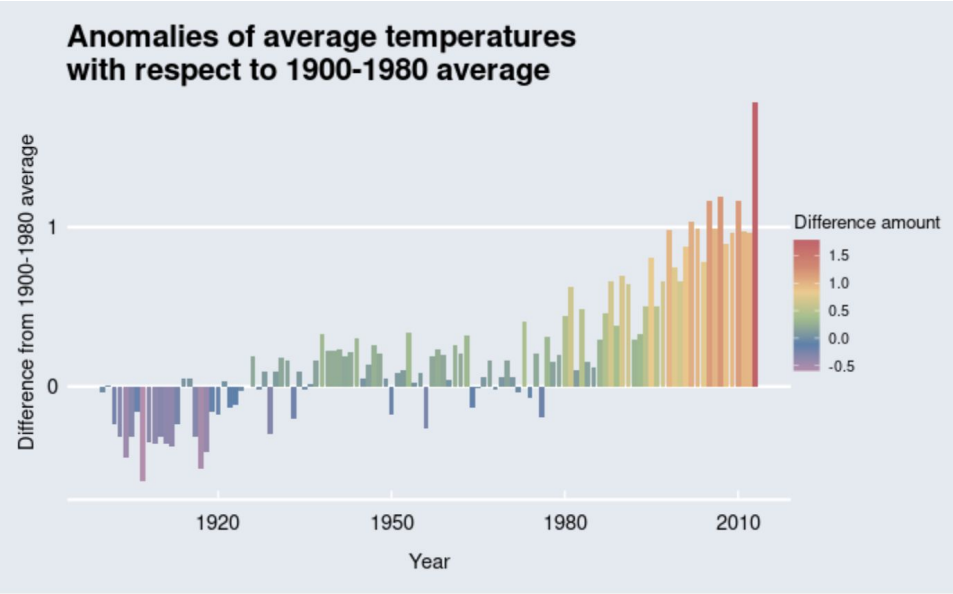


Visualizations





Hypothesis Test #1

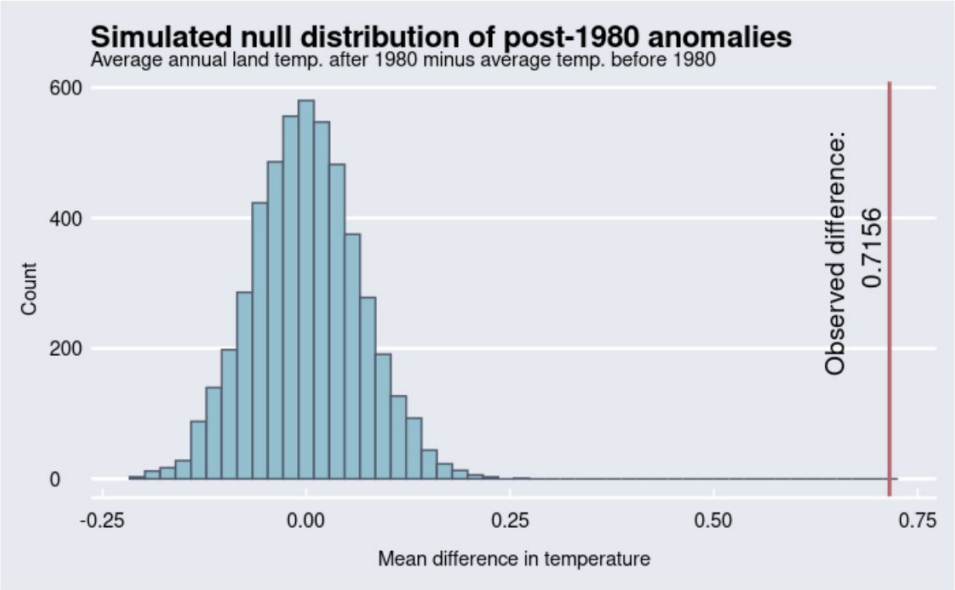


$$H_o : \mu_{post1980} \leq \mu_{pre1980}$$

$$H_a : \mu_{post1980} > \mu_{pre1980}$$

The p-value from a simulation based hypothesis test was 3.866e-14, which in context of our question, indicates that there is evidence to support a statistically significant increase in temperature over the 20th century. This is important because this is one of the main arguments used against climate change.

Hypothesis Test #1

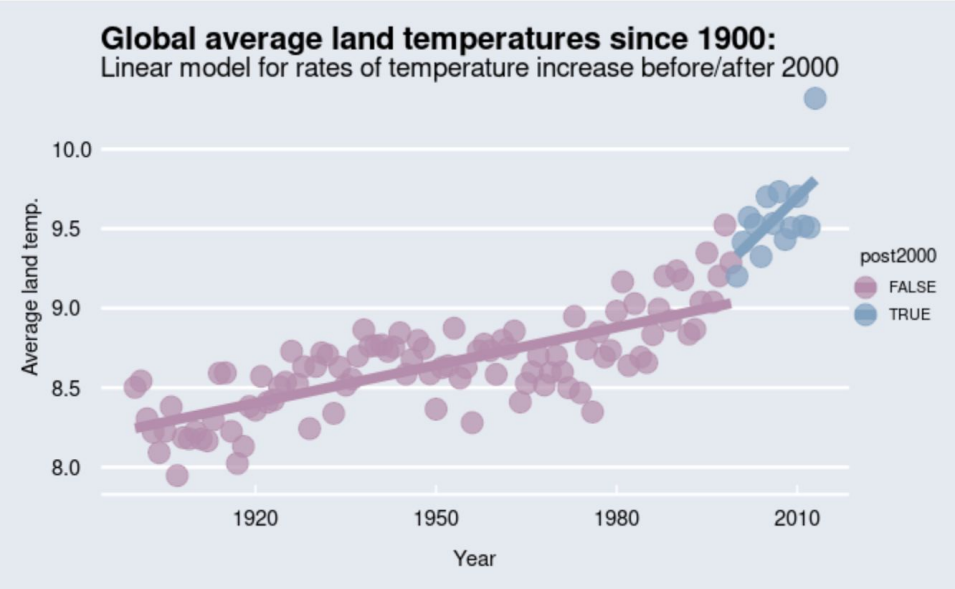


$$H_o : \mu_{post1980} \leq \mu_{pre1980}$$

$$H_a : \mu_{post1980} > \mu_{pre1980}$$

The p-value from a simulation based hypothesis test was 3.866e-14, which in context of our question, indicates that there is evidence to support a statistically significant increase in temperature over the 20th century. This is important because this is one of the main arguments used against climate change.

Hypothesis Test #2



$$H_o : \beta_{post2000} \leq \beta_{pre2000}$$

$$H_a : \beta_{post2000} > \beta_{pre2000}$$

A simulation based hypothesis test yielded a p-value = 0.0214 for the null hypothesis: that the rate of temperature change after 2000 was less than or equal to the rate of temperature change before 2000. This is statistically significant and we reject the null hypothesis, which provides evidence that the rate of temperature change has increased from years 1900-2000 to years after 2000.

The same test using years before and after 1980 found a p-value of 0.0592. This suggests we may have insufficient evidence of a significant increase in the average degree of temperature change from before 1980 to after 1980.

Discussion

For our final hypothesis test, we looked at whether the average temperature change in Europe and North America respectively were greater than the global average.

- ➔ We failed to reject the null hypothesis due to large p-values, but we did find that Europe had a much smaller confidence interval range of change in mean temperature than North America did.
- ➔ This allows us to draw connections between these two continents and their differing approaches to tackling climate change.

Overall, our analysis of the dataset found support of climate change and highlighted that it has become a more dangerous issue in recent years. It also provided evidence of the importance of political initiatives to prevent climate change.