Lagged Connections: Exploring the Spatial Distribution of Internet Access + Usage in 2021

Avery Jonas
Pratt Institute
Spatial Statistics



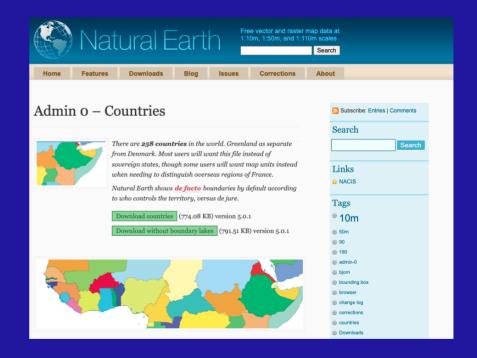
The Foundational Inquiry

- Interest in Distribution of Internet/Broadband Access + Usage Around the World (global spatial autocorrelation)
- ...by country (grid = cultural vector of nation-states)
- where in the world is there the most/least access to the internet...internet user and mobile subscribers (hotspot/coldspot analysis)

The Data



"The Inclusive Internet Index, commissioned by Facebook and developed by The Economist Intelligence Unit, seeks to measure the extent to which the Internet is not only accessible and affordable, but also relevant to all, allowing usage that enables positive social and economic outcomes at the individual and group level."



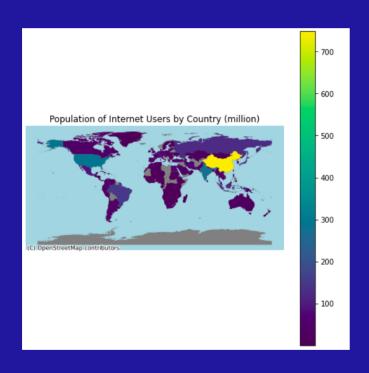
Vector files for recognized sovererign nations

Methods

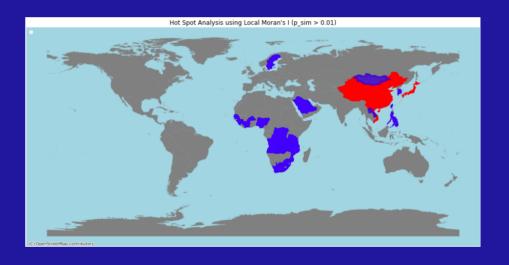
- Read in both Inclusive Index dataset and Natural Earth vector file
- · Clean and wrangle data
 - correct country names (Congo, U.S., Tanzania, Ivory Coast/Côte d'Ivoire [directly on .csv file])
 - create a 'Country' feature in geodataframe, converting 'SOVERERIGNT' type object -> string
- · drop null values in dataframe
- Merge (pandas) dataframe and geodataframe 'on' the 'Country' column

- Query for 2021 ("Edition = 'E5"")
- Point Pattern Analysis
 - Plot population of internet users by country and test for spatial autocorrelation
 - Plot percentage of population that are internet users by country and test for spatial autocorrelation
 - Plot mobile subscribers by country and test for spatial autocorrelation
- Look for other appropriate features for analysis

Population of Internet Users by Country

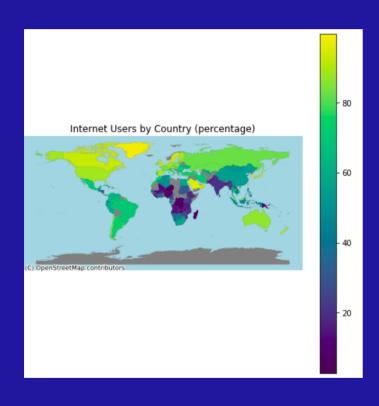


Moran's I statistic = 0.193 -> little to no global spatial autocorrelation p-value = 0.0009 -> reject null hypothesis

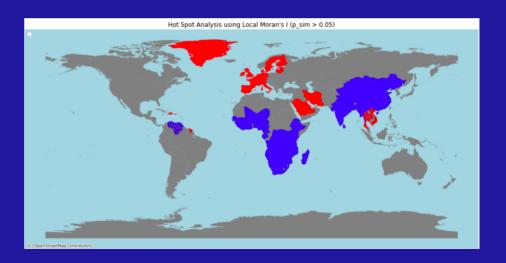


China, Vietnam and Japan are hot spots, while Sweden, Saudi Arabia and parts of West Africa and southern Africa are cold spots. Smaller East Asian countries are doughnuts.

Percentage of Internet Users by Country

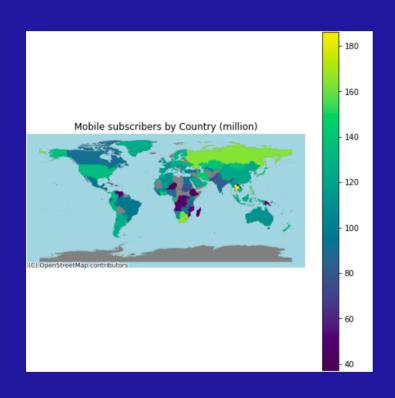


Moran's I statistic = 0.705 -> positive global spatial autocorrelation p-value = 0.0001 -> reject null hypothesis

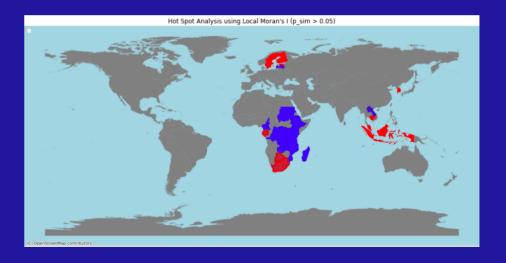


Most of Europe, Saudi Arabia, Iran, and parts of the Caribbean are hotspots. West, Southern, and East Africa as well as most of Asia are coldspots. Venezuela is a doughnut.

Mobile Subscribers by Country



Moran's I statistic = 0.23 -> very little global spatial autocorrelation p-value = 0.0001 -> reject null hypothesis



Sweden, Finland, Indonesia, Malaysia, Cambodia and South Korea are hotspots. East and Central Africa are coldspots, while Southern Africa and Gabon are diamonds. Laos is a doughnut.

Other Features of Interest

- Infrastructure
 - Average Broadband Capacity
 - Average Fixed Broadband Upload Speed
 - Average Fixed Broadband Download Speed
 - Network Coverage
 - Wireless and Broadband Operators' Market Share

*units were not provided and could not be inferred for these features, due to inability to access the Inclusive Internet Index Excel workbook

Discussion

- Effectiveness of Percentage vs. Count
 - Ignored percentage of mobile subscribers per country because in some countries the number of mobile subscribers was larger than the population
- Percentage of Internet Users per Country exhibits spatial correlation
- Finer tuned analysis regression models
- May be interesting to look at the years 2017-2021 | show the effects of COVID on internet infrastructure
- ~27 countries are not included in the Inclusive Index dataset -> noticeable or maps

THANK YOU!