**✅ Short-Term Recommendation: Focus on the United States 🇺🇸**

**Reasons:**

* **Strong data accessibility** – data is easy to access and reproduce (e.g., EPA Air Quality System, CDC asthma prevalence, BRFSS health surveys);
* **More available subnational asthma prevalence data** (e.g., at the county level), suitable for spatial statistics and predictive modeling;
* **Significant regional variation in air pollution** (e.g., California vs. the Midwest), making it easier to model spatial risk patterns;
* Well-suited for training **spatial regression models**, **GWR**, and **machine learning models**, and for conducting **policy simulations** (e.g., pre- and post–Clean Air Act analysis).

**🧪 Suggested Research Directions (U.S. Context):**

| **Type** | **Analytical Idea** |
| --- | --- |
| 🗺️ Spatial Analysis | Identify areas where asthma rates strongly align with PM2.5 levels using LISA and Moran’s I |
| ⏳ Temporal Trends | Analyze long-term correlations between air pollution and asthma hospitalization, explore lag effects and temporal patterns |
| 🤖 Machine Learning | Use XGBoost or Random Forest to predict county-level asthma risk and interpret key factors (e.g., NO₂, PM2.5, SES) |
| 📉 Policy Evaluation | Compare asthma trends before and after the Clean Air Act using DID or synthetic control methods |

**📦 Recommended U.S. Data Sources:**

| **Data Category** | **Source** | **Description** |
| --- | --- | --- |
| PM2.5, NO₂ | **EPA AQS** | Daily-level air pollution data at county/city scale |
| Asthma Prevalence | **CDC PLACES / BRFSS** | County-level prevalence rates available |
| Socioeconomic Data | **U.S. Census / ACS** | Population, income, racial composition, and more |
| Health Records | **HCUP (Healthcare Cost and Utilization Project)** | Asthma hospitalization data (some access restrictions apply) |