Methodology for Quarterly Internet Speed Measurement at Country/Admin1/Admin2 level

April 25, 2025

He Wang

This document outlines the methodology for calculating **quarterly country/admin1/admin2-level internet speed from 2019 Q1 to 2024 Q4**, using global fixed and mobile network performance data. It covers data sources, versioning, processing, and output variables to ensure transparency and reproducibility.

1. Data Source

The calculation of internet speed relies on two primary datasets:

1.1 Internet Speed Data

- Source: Speedtest by Ookla – Global Fixed and Mobile Network Performance Maps
- Description: This dataset provides global fixed broadband and mobile (cellular) network performance metrics in zoom level 16 Web Mercator tiles (approximately 610m by 610m at the equator). The data is updated quarterly. On average, each quarter includes 4.9 million tiles for fixed internet data and 3.3 million tiles for mobile internet data. The data is aggregated from real-world user tests conducted via Ookla's Speedtest application.
- Access Link: Ookla Open Data

1.2 World Map Data

- Source: GADM world map, version 4.1
- Description: This dataset provides high-definition administrative boundaries for countries and subnational units, used for spatial aggregation and analysis of internet speed data at the national level.
- Access Link: GADM World Map

2. Version and Date

- **Ookla Data:** Downloaded using the *ooklaOpenDataR* R package on March 8, 2025.
- **World Bank Map Data:** Accessed on March 8, 2025, using the version labeled "Last Updated: Mar 19, 2020".

3. Procedures

The original datasets provide average download/upload internet speed (kbps), latency (ms),[1] number of tests, and number of devices performing the tests at the zoom-16 tile level for each quarter. The following steps outline the methodology for processing and aggregating the data to generate quarterly country-level internet speed estimates.

Understanding the Quadkey System

The dataset uses a quadkey system, a spatial index that hierarchically partitions the Earth's surface into tiles based on a Web Mercator projection (EPSG:3857).

- Tile Hierarchy: At zoom level 0, the entire world is represented by a single tile. Each zoom level subdivides every tile into four smaller tiles, exponentially increasing spatial resolution.
- Quadkey Encoding: Each tile is uniquely identified by a quadkey, a string of digits where each digit (0,1,2,3) represents the tile's position in a parent tile's subdivision.
- Zoom-16 vs. Zoom-14: At zoom-16, tiles are 610.8m × 610.8m at the equator, while zoom-14 tiles cover approximately 2.5 km × 2.5 km, providing a coarser but computationally efficient resolution.

Example: Quadkey '123' at Zoom Level 3

A quadkey is read digit by digit, where each digit (0-3) determines the tile's position at a given zoom level:

- '1' → Selects the upper-right tile at zoom-1.
- '2' → Chooses the lower-left tile within that upper-right tile.
- '3' → Picks the lower-right tile within the lower-left tile.

Step 1: Adjusting Spatial Granularity

The raw dataset is provided at zoom-16, meaning each tile represents a very small area (~610.8m × 610.8m). To improve computational efficiency and ensure meaningful aggregation, I reduced the granularity to zoom-14, where each tile represents approximately 2.5 km × 2.5 km.

Step 2: Computing Tile Centroids

Each zoom-14 tile is identified by a 14-digit quadkey, and its centroid coordinates (latitude, longitude) are computed as follows:

1. Decode the quadkey → Convert the quadkey into tile coordinates (X, Y) and zoom level (Z) using base-4 encoding.

[1] The average download speed, upload speed, and latency of all *tests* performed in the tile, represented in kilobits per second or millisecond.

2. Convert to Latitude/Longitude → Compute the centroid in geographic coordinates using:

$$\text{Longitude} = \left(\frac{X + 0.5}{2^Z} \times 360.0\right) - 180.0$$

$$\text{Latitude} = \text{atan}\left(\sinh\left(\pi \times \left(1 - \frac{2 \times (Y + 0.5)}{2^Z}\right)\right)\right) \times \frac{180}{\pi}$$

Step 3: Assigning Internet Speed Data to Countries

Using the World Bank Official Boundaries dataset, I performed a spatial lookup to determine which country each zoom-14 tile centroid belongs to.

Step 4: Aggregating Internet Speed at the Country/Admin1/Admin2 Level

For each country/admin1/admin2 unit,

- Weighted average internet speed and latency are computed for mobile and fixed internet separately, with the number of *tests* as weights:

$$\text{Country Speed} = \frac{\sum(\text{Speed} \times \text{Tests})}{\sum \text{Tests}}$$

- Similarly, weighted median internet speed and latency are calculated for mobile and fixed internet separately, with the number of *devices* as weights. Detailed procedures are as follows:
    1. The tiles within each country-year-quarter are first sorted by the variable of interest (internet speed or latency).
    2. The cumulative sum of devices performing tests is then calculated.
    3. The tile where the cumulative sum reaches the 50th percentile is identified, and the internet speed or latency in this tile represents the weighted median.

4. Output Variables

The final dataset provides quarterly country/admin1/admin2-level estimates of internet speed, including download/upload speeds, latency, and test volumes for both mobile and fixed internet connections. The table below summarizes the key output variables and their meanings.

Variable Name	Description
country_code	Three-letter country code
country_name	Country name
admin1_code	Admin-1 level unit code
admin1_name	Admin-1 level unit name
admin2_code	Admin-2 level unit code
admin2_name	Admin-2 level unit name
year	Year of observation
quarter	Quarter of observation (1 - 4)
mobile_upload_speed	Average mobile internet upload speed (kbps)

mobile_download_speed	Average mobile internet download speed (kbps)
mobile_latency	Average mobile internet latency (ms)
mobile_median_upload_speed	Median mobile internet upload speed (kbps)
mobile_median_download_speed	Median mobile internet download speed (kbps)
mobile_median_latency	Median mobile internet latency (ms)
mobile_tests	Total number of tests for mobile internet
mobile_devices	Total number of devices performing mobile speed tests
fixed_upload_speed	Average fixed internet upload speed (kbps)
fixed_download_speed	Average fixed internet download speed (kbps)
fixed_latency	Average fixed internet latency (ms)
fixed_median_upload_speed	Median fixed internet upload speed (kbps)
fixed_median_download_speed	Median fixed internet download speed (kbps)
fixed_median_latency	Median fixed internet latency (ms)
fixed_tests	Total number of tests for fixed internet
fixed_devices	Total number of devices performing fixed speed tests

Contact Information

- **Name:** He Wang
- **Email:** hwang21@worldbank.org