



Regional Government Success Scorecard

Methodology Note





Content

1	Overview of the Index	03
2	Data Sources	07
3	Methodology	10

1.1 Purpose

The Regional Government Success Scorecard (RGSS) measures the performance of sub-national governments (e.g. at the regency (*kabupaten*) and city (*kota*) levels for Indonesia), in a way that is comparable across regions and fair to their structural starting points. It has two objectives:

Objective 1:

Measuring regional governance success. Three transparent pillar scores and an overall RGSS score together capture how well-resourced each local government is, how capable it is, and what performance it delivers for its citizens, adjusting for the underlying Foundational Environment (FE).

Objective 2:

Identifying factors behind successful performances. E.g., Analysis of correlations between Performance, Inputs, and Capabilities pillars identifies the main drivers of success, giving senior stakeholders evidence-based reference points for development priorities.

1.2 Framework

The RGSS organises every dimension within one of four components, three of which are measured pillars:

- **Pillar 1, Inputs (I):** resources available to the government
- **Pillar 2, Capabilities (C):** institutional capacity and quality
- **Pillar 3, Performance (P):** citizen-facing results delivered by the government. It contains a combination of output dimensions and outcome dimensions
- **Foundational Environment (FE):** structural conditions largely outside the government's direct control

Inputs and Capabilities together form ‘Local Government Resource and Capabilities’, while Performance captures citizen-facing results. Foundational Environment is not scored; it serves as a context-adjustment layer. The adjustment operates at the dimension level for simplicity: for every dimension within Inputs, Capabilities and Performance, each region’s score is compared against the scores of structurally similar peers, so that local governments are assessed in context rather than ranked on raw scores alone. The adjusted scores are the same as if they were adjusted at indicator-metric level.

The methodology for adjustment is further explained in **section 3.3**.

Each pillar is made up of a number of dimensions. Each score is the simple average of its normalised indicator-metric scores. Each pillar score is the simple average of its adjusted dimension scores. The overall RGSS score is the simple average of the three pillar scores: $(\text{Capabilities} + \text{Inputs} + \text{Performance}) / 3$.

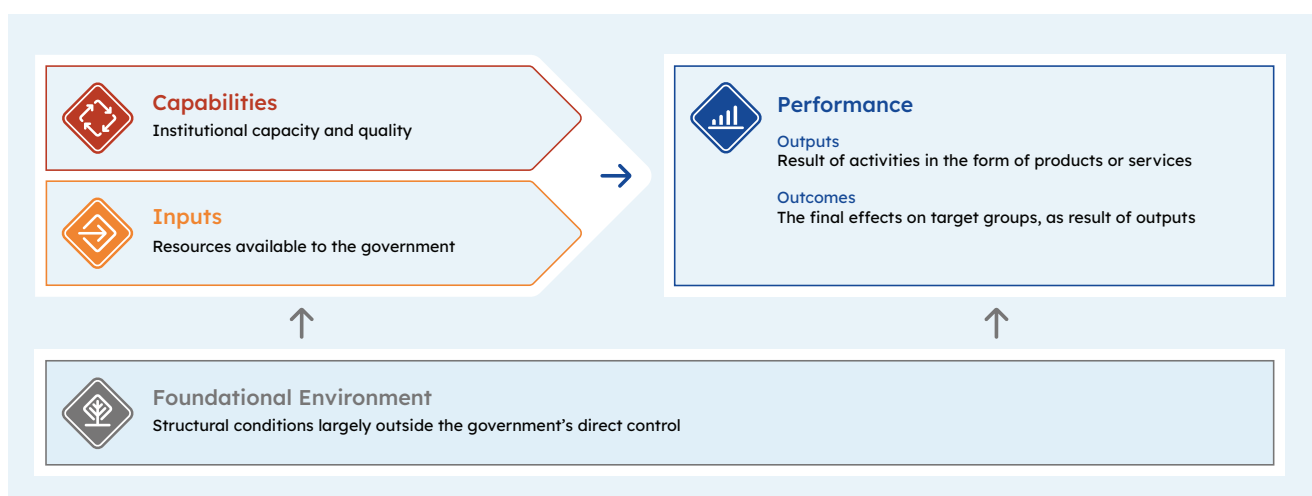


Figure 1: RGSS Framework

Each region thus receives adjusted dimension scores, three pillar scores, and an overall RGSS score, each accompanied by contextual notes to support interpretation.

1.3 Notes on the framework

Two points on the framework are worth noting:

Performance is a combination of output and outcome.

We use this specific architectural choice because (1) Outputs are sometimes recognised as intermediate Outcomes in the literature; (2) At a multi-sector government level, what are considered Outcomes to one government agency may be treated as Outputs to another; and (3) Specific data availability in local contexts may not give a balanced view if they are weighted equally. We make this specific conceptual (not computational) distinction within the performance pillar. This preserves the RGSS's ability to track the full input-output-outcome chain in contexts where data is available.

GDP per capita sits outside the headline score.

GDP per capita is kept as a separate analytical lens alongside population, special-autonomy status, and archipelagic status, rather than feeding into the headline calculation. Treating it as a contextual adjustment would risk adjusting away the very performance the Scorecard is measuring.

1.4 Contextual Data and Filters

The RGSS publishes a set of contextual data points alongside its scored indicators. These are not part of the scorecard calculation and do not affect scores, but they allow users to filter rankings and compare within sub-groups that share governance characteristics beyond what DPC captures. The relevant classifications vary by country; the Indonesia pilot uses two.

Archipelagic status. Regions whose territory is distributed across islands face service delivery costs and connectivity constraints that contiguous land-based regions do not. In Indonesia, archipelagic classification follows the criteria set out in the '*2017 Draft Law on Archipelagic Regions*', which designates provinces, and regencies and cities where sea area exceeds land area as archipelagic (greater than 50% of total area). Filtering by archipelagic status allows users to compare regions within this sub-group on more equal footing.

Special autonomy and fiscal decentralisation status. Some regions operate under asymmetric governance arrangements that alter their administrative authority and revenue base relative to standard regions. This affects the fiscal environment of regencies (kabupaten) and cities (kota) within those provinces, even though the special status is held at provincial level and does not apply directly to sub-provincial units. The RGSS records this as a contextual variable to support interpretation, not as a scoring input.

In future country applications, other contextually relevant classifications - such as conflict-affected status, climate vulnerability designations, or sub-national administrative tier - may be added as filter variables depending on the country's governance landscape. Characteristics that affect the governance environment but are not fully captured by the scored Foundational Environment dimensions can be surfaced as filters without distorting the scores.

2

Data Sources

2.1 Data Sources

Every dimension in the framework above is operationalised through one or more indicator-metrics, each drawn from an identified Indonesian government data source. The list appears below, organised by pillar, dimension, and indicator-metric in descending order.

Pillars	Dimensions	Indicator-metric	Data Source
 Capabilities	Technology & Innovation	Use of ICT in government service delivery	Ministry of Administrative and Bureaucratic Reform (Kementerian PANRB)
		Regional innovation index	The Ministry of Home Affairs
	Accountability	Integrity score for anti corruption efforts	Corruption Eradication Commission (KPK)
	Financial Management	Regional fiscal capacity index score	Ministry of Finance
		Regional fiscal independence index score	Ministry of Finance
		Budget Execution rate	Ministry of Finance
		Quality of public procurement	National Public Procurement Agency (LKPP)

Pillars	Dimensions	Indicator-metric	Data Source
 Capabilities	Public Service	Service quality in investment and revenue offices	Ministry of Administrative and Bureaucratic Reform (Kementerian PANRB)
 Inputs	Financial Resources	Local budget amount	Ministry of Finance
	Human Resources	Number of civil servants	National Civil Service Agency
 Performance	Attracting Investment (Output)	Growth of domestic investment	Ministry of Investment and Downstream Industry/Investment Coordinating Board (BKPM)
		Growth of foreign investment	Ministry of Investment and Downstream Industry/Investment Coordinating Board (BKPM)
	Economic Prosperity	Adjusted real annual expenditure per capita	BPS-Statistics Indonesia
		Economic growth rate	BPS-Statistics Indonesia
	Employment	Unemployment rate	BPS-Statistics Indonesia
		Employment opportunity	BPS-Statistics Indonesia
		Labour productivity	BPS-Statistics Indonesia

Pillars	Dimensions	Indicator-metric	Data Source
 Performance	Income & Poverty	Share of population in poverty	BPS-Statistics Indonesia
		Poverty gap index	BPS-Statistics Indonesia
		Poverty severity index	BPS-Statistics Indonesia
		Gini coefficient/ratio	BPS-Statistics Indonesia
	Education	Participation in formal education (average years of schooling)	BPS-Statistics Indonesia
		Participation in formal education (school years expectancy)	BPS-Statistics Indonesia
	Health	Life expectancy at birth	BPS-Statistics Indonesia
	Gender Equality	Active participation of women in politics & economy	BPS-Statistics Indonesia
		Gender equality in human development	BPS-Statistics Indonesia
	Environmental Quality	Environmental quality	BPS-Statistics Indonesia
 Foundational Environment	Natural Resources	Ratio of natural resource revenue sharing to regional revenue	Ministry of Finance
	Geographic Condition	Construction cost index	BPS-Statistics Indonesia
		Disaster risk index score	The National Agency for Disaster Countermeasure (BNPB)

Note: the pilot uses the most recent year of data available from each source (predominantly 2024).

The methodology converts raw data into comparable, context-adjusted scores in four steps, using Dynamic Peer Comparison (DPC) as the context-adjustment mechanism.

3.1 Scoring Pipeline

Each region's score is built in four steps:

Step

- 1 Dimension scores.** Each indicator-metric is first cleaned (imputation for missing values; winsorisation at the 1st and 99th percentiles to limit the influence of outliers) and then rescaled to a common 0 to 1 range using min-max normalisation. Each dimension score is the simple average of its normalised indicator-metrics. The adjusted scores are the same as if they were adjusted at indicator-metric level, and are done at the dimension level for simplicity.
- 2 Context adjustment (DPC).** Each dimension score is adjusted using Dynamic Peer Comparison. For every dimension, we subtract the peer average, which is the average score of all other local governments in the same class (region), weighted so that region with more similar Foundational Environment profiles receive greater weight.
- 3 Pillar scores.** Each pillar score (Capabilities, Inputs, Performance) is the simple average of its adjusted dimension scores.
- 4 Overall score.** The overall RGSS score is the simple average of the three pillar scores: **(Capabilities + Inputs + Performance) / 3**.

Directionality:

A positive adjusted dimension score means the region outperforms its structural peers on that dimension; a negative value means it falls short.

3.2 Adjusting Dimension Scores using Dynamic Peer Comparison (DPC)

The context adjustment uses Dynamic Peer Comparison (DPC), which compares each local government to others in the same class rather than to a single fixed group, giving greater weight to regions that are more similar in their Foundational Environment (FE) profile.

Structural similarity between any two local governments is measured by the Euclidean distance between them in FE space. The two FE dimensions, Geographic Condition and Natural Resources, are normalised to the same 0 to 1 scale so that each contributes equally to the distance calculation. This adjustment methodology holds for one or more FE dimensions. Any two local governments with similar FE profiles sit close together, while two with more different profiles are further apart. The weight given to each peer is the inverse of this distance:

To state this precisely: let 'i' denote the focal region, and let 'j' index all other regions within the same class. GC_i and NR_i denote region i's normalised Geographic Condition and Natural Resources scores (and similarly GC_j and NR_j for region j), and s_j denotes region j's dimension score. Note that in Indonesia, regencies (*kabupaten*) and cities (*kota*) are treated as separate classes: all distances and peer averages are computed within each class, so a Kabupaten is never compared to a city, and vice versa.

The Euclidean distance between 'i' and 'j' in FE space is:

$$d_{ij} = \sqrt{(GC_i - GC_j)^2 + (NR_i - NR_j)^2} \quad (1)$$

The formula generalises naturally: if a future edition introduces additional FE dimensions, each extra dimension adds one more squared term inside the square root. To prevent a peer with a near-identical FE profile from dominating the peer average, distances are floored at a minimum value $\epsilon = 0.02$ (on the normalised 0 to 1 scale):

$$\tilde{w}_{ij} = \frac{1}{\max(d_{ij}, \epsilon)} \quad (2)$$

These raw weights are rescaled to sum to 1 across all peers:

$$w_{ij} = \frac{\tilde{w}_{ij}}{\sum_j \tilde{w}_{ij}} \quad (3)$$

As illustrated in Figure 2, the closer a peer is located relative to the focal region, the higher its weight becomes.

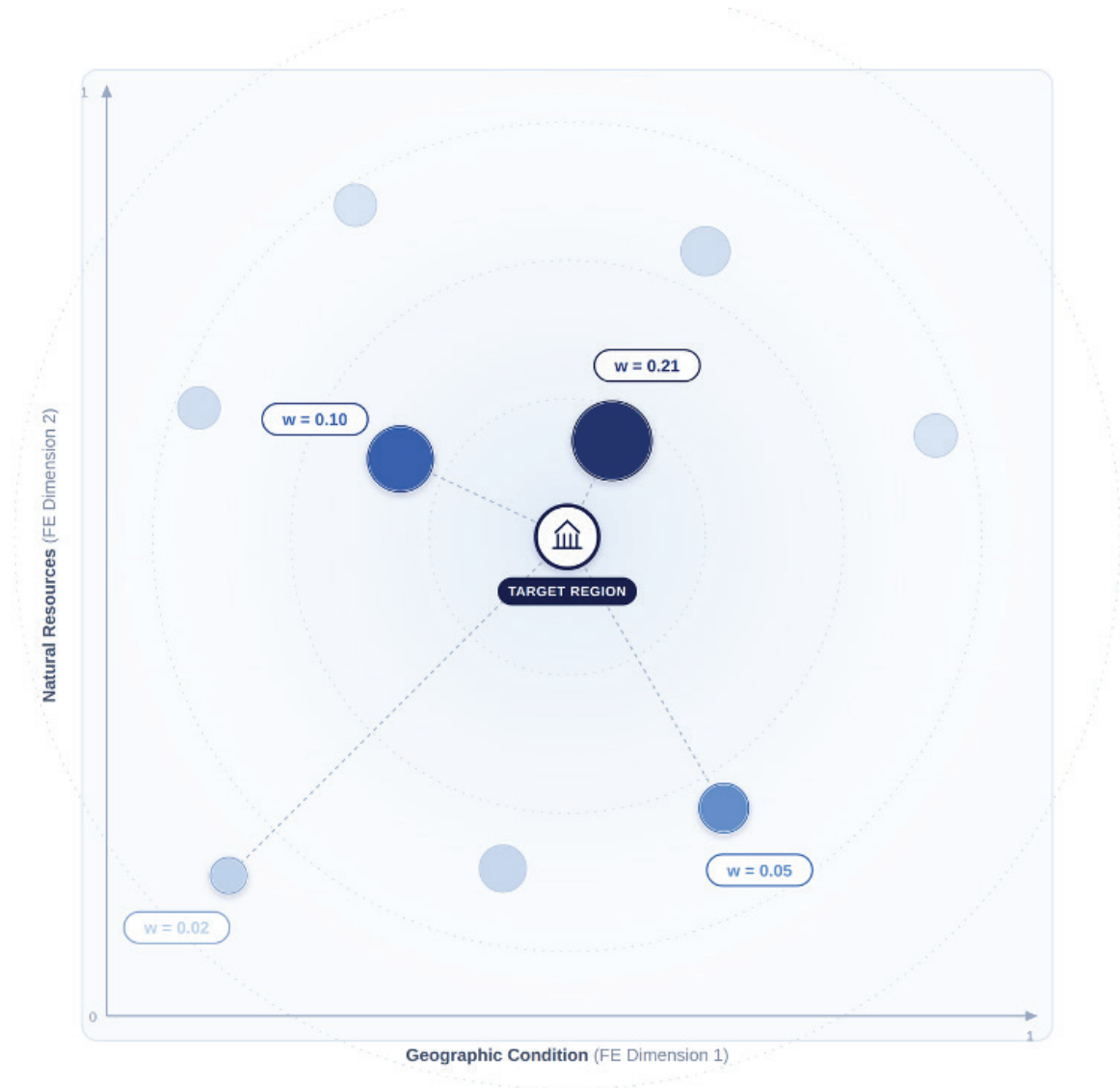


Figure 2: Dynamic Peer Comparison in Foundational Environment space
(hypothetical illustration)

Each circle is a peer region of the target region within the same class, with its position set by the two FE dimensions and its size by the peer weight, w . Regions structurally closer to the target carry more weight.

The Peer Average Dimension Score for region i is the weighted mean of all other regions' dimension scores:

$$\text{Peer Average Dimension Score}_i = \sum_j w_{ij} \cdot s_j \quad (4)$$

The adjusted dimension score for region i is the local government's actual score minus the inverse-distance-weighted peer average for that dimension (closer local governments have higher weights):

$$\text{Adjusted Dimension Score}_i = \text{Dimension Score}_i - \text{Peer Average Dimension Score}_i \quad (5)$$

After the DPC adjustment at the dimension level, each pillar score is the simple average of its adjusted dimension scores.

3.3 Why Dynamic Peer Comparison (DPC)

Alternatives were considered before adopting DPC.

Regression-based adjustment estimates the statistical relationship between FE scores and other-pillar scores, then subtracts the estimated contribution from each region's score. Although this is a common approach, regression-based methods may sometimes have sensitive regression coefficients depending on how the independent variables are entered into the equation. This statistical noise is amplified by the availability of data, multi-collinearity and measurement errors from individual datasets. The regression method also requires an explicit specification of the direction of each FE dimension's effect, such as whether natural resources are a curse or boon. Such judgements are often context-specific instead of being universal in nature.

Fixed banding groups regions into discrete FE categories and compares them within each group. This introduces two problems: (1) thresholds are intrinsically arbitrary, and can assign structurally similar regions to different bands; (2) the number of bands grows unwieldy when a higher number of FE variables are used.

DPC allows us to focus on improving comparability between local governments as needed in the scorecard while avoiding some of the challenges above. DPC does not require assumptions about the direction or magnitude of FE's effect, uses a continuous rather than discrete-thresholds, and is also compatible where the FE has more dimensions.

3.4 Step-by-step summary

- **Imputation.** Fill missing values at the indicator-metric level. We adopted a simple imputation method by a linear regression between the indicator-metric and GDP per capita, within each class (region).
- **Outlier treatment.** Winsorise indicator-metrics at the 1st and 99th percentiles for outliers.
- **Normalisation.** Rescale each indicator-metric to a 0 to 1 range using min-max normalisation.
- **Dimension scores.** Each dimension score is the simple average of its normalised indicator-metrics.
- **FE distance.** Compute the Euclidean distance between every pair of regions in FE space, within each class (region).
- **Peer averages.** For each dimension, compute the inverse-distance-weighted average of all other local governments' dimension scores within the same class.
- **Adjusted dimension scores.** Each adjusted dimension score is the raw dimension score minus its peer average using DPC.
- **Pillar scores.** Simple average of the adjusted dimension scores within each pillar (Capabilities, Inputs, Performance).
- **Overall scores.** Simple average of the three pillar scores.
- **Peer identification and ranking.** Identify each region's closest structural peers and compute its rank within that peer set.