

Task 1 — Global Population Estimate & Projections (Power BI Dashboard)

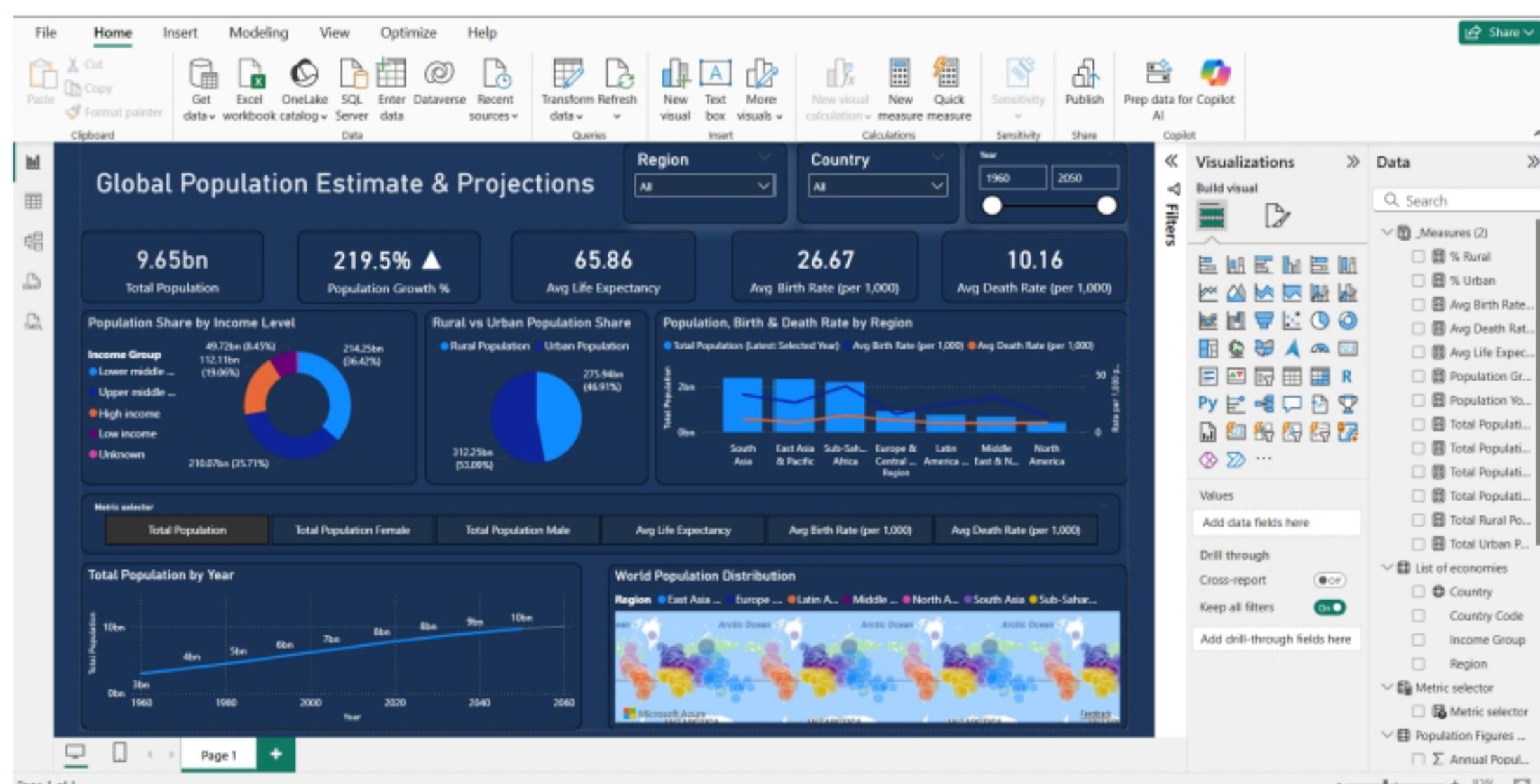
1. Scenario Summary

The chosen scenario is the **Global Population Estimate & Projections** dataset combining historical figures (1960–2022) and forward projections (up to 2050) for all recognised countries. Core indicators include **total population** (female/male), **urban and rural population**, **life expectancy**, and **birth/death rates**. A supplementary **Country Groupings** file classifies each country into a **Region** (e.g., East Asia & Pacific) and **Income Group** (Low, Lower middle, Upper middle, High), using World Bank conventions.

Why this matters: population dynamics influence almost every long-horizon decision—food security, climate and resource planning, health care capacity, education, housing, labour markets, and infrastructure. A dashboard that lets a user move from the global picture down to regions and countries, with **time as a first-class dimension**, supports exploration of demographic transition patterns (declining fertility, increasing longevity, urbanisation) and reveals how different regions are at different stages of that transition.

Dashboard intent: present an **interactive, single-page** analytical overview that:

- shows long-term **trends** (1960–2050),
- compares **regions** and **income groups**,
- surfaces **urbanisation** and **mortality/fertility** patterns,
- and enables **country-level** exploration when needed.



2. Initial Exploration & Specific Objectives

2.1 Exploratory observations (Power Query/Data View)

- **Temporal continuity:** complete yearly sequence 1960–2050 enables line charts and forecasting.
 - **Missingness pattern:** life expectancy and rate indicators show patchy coverage in early decades for some countries; population counts are largely complete.
 - **Duplicates:** none for the natural key (**Country Code, Year**).
 - **Type sanity:** numeric indicators stored as numbers; time as integer year.
 - **Regional heterogeneity:** early pivots showed rapid growth in **South Asia** and **Sub-Saharan Africa**, with **Europe & Central Asia** much flatter or declining in later years.
 - **Urbanisation:** global **urban share > rural** after the mid-2000s; still varies by income group.
-

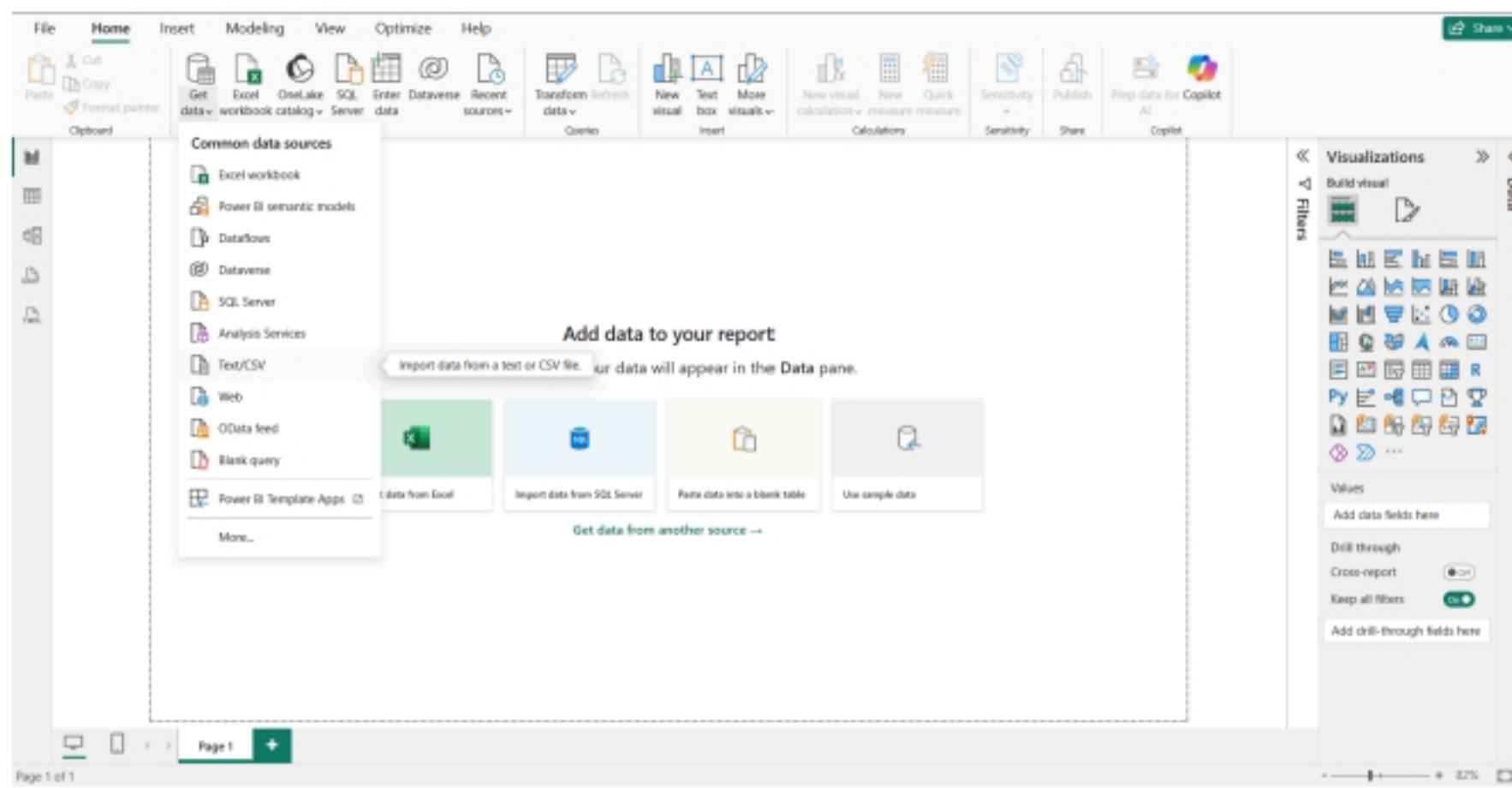
2.2 Objectives defined from EDA

1. **Trend objective:** show global and region/country **population trajectory** 1960–2050 with simple access to growth.
 2. **Structure objective:** explain **composition**—by **Income Group** and **Rural vs Urban**—with part-to-whole visuals that are compact and legible.
 3. **Comparative objective:** juxtapose **Population** with **Birth/Death rates** at the **region** level to communicate demographic transition.
 4. **Exploration objective:** provide **slicers** for **Region, Country, and Year range**; add a **metric selector** to switch the main line chart between Total Population, Female/Male, Life Expectancy, Growth %, etc.
 5. **Design objective:** apply a **high-contrast dark theme** for readability and focus, keep clutter minimal, and conform to best-practice layout (overview → comparisons → trends → geography).
-

3. Data Preparation, Modelling & DAX

3.1 Data preparation

1. **Import**
 - *Get data* → *Text/CSV* for *Population Figures & Projections.csv*.
 - *Get data* → *Excel* for *Country Groupings.xlsx*.



2. Data types

- Year → Whole number; Country Code/Name → Text; indicators → Decimal Number.

3. Text cleaning

- Trim and Clean on Country and Country Code to prevent join errors.

4. Column simplification

- Rename long indicator headers to user-friendly names (e.g., *Population Female*, *Population Male*, *Birth rate (per 1,000)*, *Death rate (per 1,000)*).

5. Missing values

- Left nulls for rates/life expectancy where gaps are long (honest representation).
- One missing *Income Group* recoded to “Unknown”.

The screenshot shows the Power BI 'Transform' ribbon with various data manipulation tools like Transpose, Replace Values, Unpivot Column, etc. A table is open with columns: Country Code, Country, Region, Income Group. A 'Replace Values' dialog box is overlaid on the table. In the dialog, 'Value To Find' is set to 'null' and 'Replace With' is set to 'Unknown'. The table data includes rows for Georgia, Germany, Ghana, Gibraltar, Greece, Grenada, Guam, Guatemala, Guinea, Guyana, Haiti, Honduras, Hong Kong, Hungary, Iceland, India, Indonesia, Iran, Iraq, Ireland, Isle of Man, Israel, Italy, Jamaica, Japan, and Jordan.

6. Validation

- Checked duplicates on (Country Code, Year) → none; ensured year range 1960–2050 consistent across countries.
-

3.2 Star-schema model

- **Fact:** Population Figures & Projections (one row per (Country Code, Year) with indicators).
- **Dimensions:**
 - List of economies (Country Code, Country, Region, Income Group).
 - Year (generated) to enable time intelligence.
 - Metric selector (field parameter table).
 - _Measures (empty physical table to hold measures for organisation).

Relationships

- List of economies[Country Code] → Population[Country Code] (1:).
- Year[Year] → Population[Year] (1:).
- Cross-filter direction **Single** to keep propagation predictable and fast.

The screenshot shows the Power BI Data Model view. On the left, there's a diagram of data relationships between tables: 'List of economies' (Country, Country Code, Income Group, Region) connects to 'Population Figures & Projections' (Annual Population Growth(Official), Birth rate, Country, Country Code, Death rate, Life expectancy, Population Female, Population Male, Rural population, Time Code, Total Population, Urban population, Year) via a '1' to '1' relationship. This table also connects to the 'Year' table (Date, Year) via a '1' to '1' relationship. The 'Year' table connects to the same 'Population Figures & Projections' table via another '1' to '1' relationship. On the right, the 'Properties' pane is open for the 'Year' table, showing settings for 'Cards', 'Metric selector', and 'Pin related fields to top of card'. The 'Data' pane lists various population metrics like 'Total Rural Population', 'Total Urban Population', etc.

3.3 Calendar (Year) table

The screenshot shows the Power BI Data Model view. On the left, the DAX code for generating a calendar table is displayed:

```

1 Year =
2 VAR StartYear = 1960
3 VAR EndYear = 2050
4 RETURN
5 ADDCOLUMNS(
6     GENERATESERIES(StartYear, EndYear, 1),
7     "Date", DATE([Value], 12, 31)
8 )
9

```

This code creates a 'Year' table with a single column 'Date' representing dates from 1960 to 2050. Below the code, the 'Year' table is shown in the data model diagram, connected to the 'Population Figures & Projections' table. On the right, the 'Properties' pane for the 'Year' table is open, showing details like 'Name: Year', 'Description: Enter a description', 'Synonyms: year', 'Display folder: Enter the display folder', 'Is hidden: No', and 'Data type: Whole number'.

Marked as **Date table** using the Date column.

3.4 Measures (DAX) used in visuals



Totals & components

- 1 Total Population = `SUM('Population Figures & Projections'[Population Female]) + SUM('Population Figures & Projections'[Population Male])`
- 1 Total Population (Latest Selected Year) = `CALCULATE([Total Population], LASTDATE('Year'[Date]))`
- 1 Total Population Female = `SUM('Population Figures & Projections'[Population Female])`
- 1 Total Population Male = `SUM('Population Figures & Projections'[Population Male])`
- 1 Total Rural Population = `SUM('Population Figures & Projections'[Rural Population])`
- 1 Total Urban Population = `SUM('Population Figures & Projections'[Urban population])`

Composition

- 1 % Urban = `DIVIDE([Total Urban Population], [Total Population])`
- 1 % Rural = `DIVIDE([Total Rural Population], [Total Population])`

Rates / expectations

- 1 Avg Life Expectancy = `AVERAGE('Population Figures & Projections'[Life expectancy])`
- 1 Avg Birth Rate (per 1,000) = `AVERAGE('Population Figures & Projections'[Birth rate])`
- 1 Avg Death Rate (per 1,000) = `AVERAGE('Population Figures & Projections'[Death rate])`

Time-based change

```

1 Population Growth (Selection) % =
2 VAR MinY =
3 | MINX(ALLSELECTED('Year'[Year]), 'Year'[Year])
4 VAR MaxY =
5 | MAXX(ALLSELECTED('Year'[Year]), 'Year'[Year])
6 VAR StartPop =
7 | CALCULATE([Total Population], TREATAS({MinY}, 'Year'[Year]))
8 VAR EndPop =
9 | CALCULATE([Total Population], TREATAS({MaxY}, 'Year'[Year]))
10 VAR Change =
11 | IF(
12 | | MinY = MaxY,
13 | | 0,
14 | | DIVIDE(EndPop - StartPop, StartPop)
15 | )
16
17 VAR PositiveIcon = UNICHAR(9650)
18 VAR NegativeIcon = UNICHAR(9660)
19 VAR NeutralIcon = UNICHAR(8212)
20
21 RETURN
22 IF(
23 | ISBLANK(Change),
24 | BLANK(),
25 | FORMAT(Change, "0.0% ") &
26 | | IF(
27 | | | Change > 0, PositiveIcon,
28 | | | IF(Change < 0, NegativeIcon, NeutralIcon)
29 | )
30 )

```

```

1 Population YoY % =
2 VAR Prev =
3 | CALCULATE([Total Population], DATEADD('Year'[Date], -1, YEAR))
4 VAR Change =
5 | DIVIDE([Total Population] - Prev, Prev)
6 VAR PositiveIcon = UNICHAR(9650)
7 VAR NegativeIcon = UNICHAR(9660)
8 VAR NeutralIcon = UNICHAR(8212)
9 RETURN
10 IF(
11 | ISBLANK(Change),
12 | BLANK(),
13 | FORMAT(Change, "0.0% ") &
14 | | IF(Change > 0, PositiveIcon,
15 | | | IF(Change < 0, NegativeIcon, NeutralIcon))
16 )
17

```

4. Visualisations: selection & justification

This section ties each visual to **perceptual theory** and **best practice**.

4.1 KPI cards



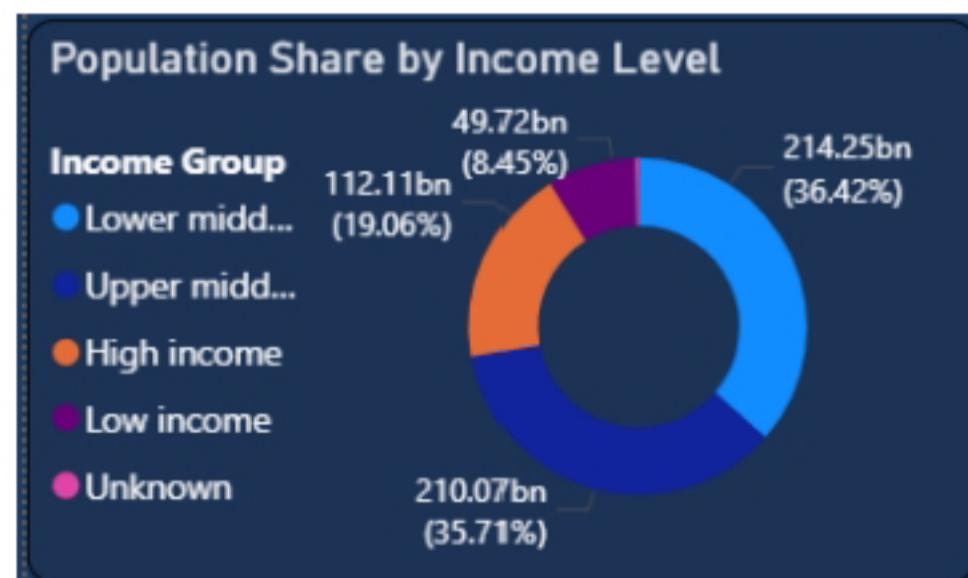
What they show:

- Total Population (latest year under current slice), Population Growth %, Life Expectancy, Birth/Death rates.

Why this is right:

- **Scorecard design:** cards deliver a quick, at-a-glance status with minimal cognitive load.
 - **Pre-attentive processing:** large numerals and consistent placement help rapid scanning.
 - Conditional icons (▲ for growth) leverage **preattentive color/orientation** to encode meaning instantly.
-

4.2 Donut chart — Population share by Income Group

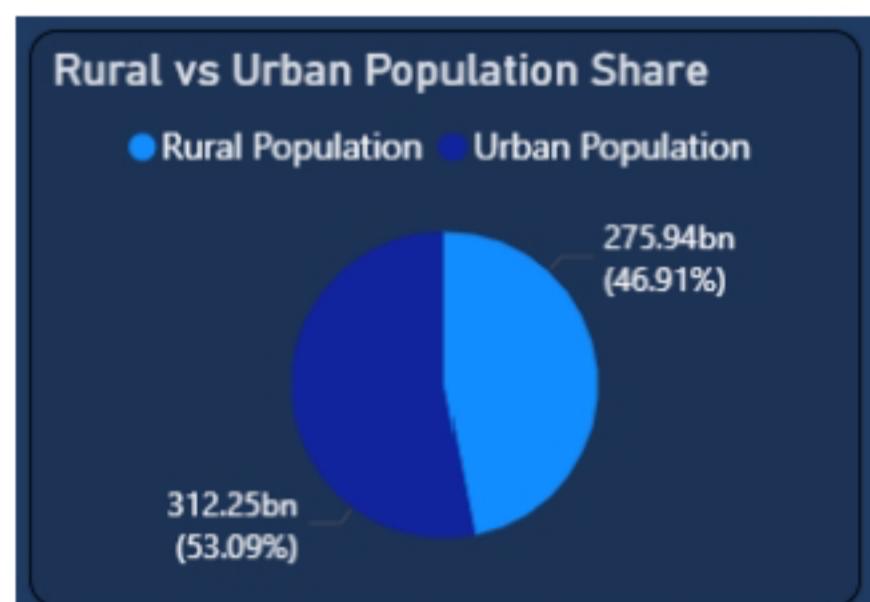


Purpose: communicate **part-to-whole** composition across four/five categories.

Theory:

- **Cleveland & McGill** caution against many slices; here the set ≤ 5 , making angle/area judgments acceptable.
 - **Legend proximity** and consistent colors reduce cognitive load (Gestalt: proximity & similarity).
 - Adds absolute value in tooltip, percentage in label (dual encoding without clutter).
-

4.3 Pie chart — Rural vs Urban share

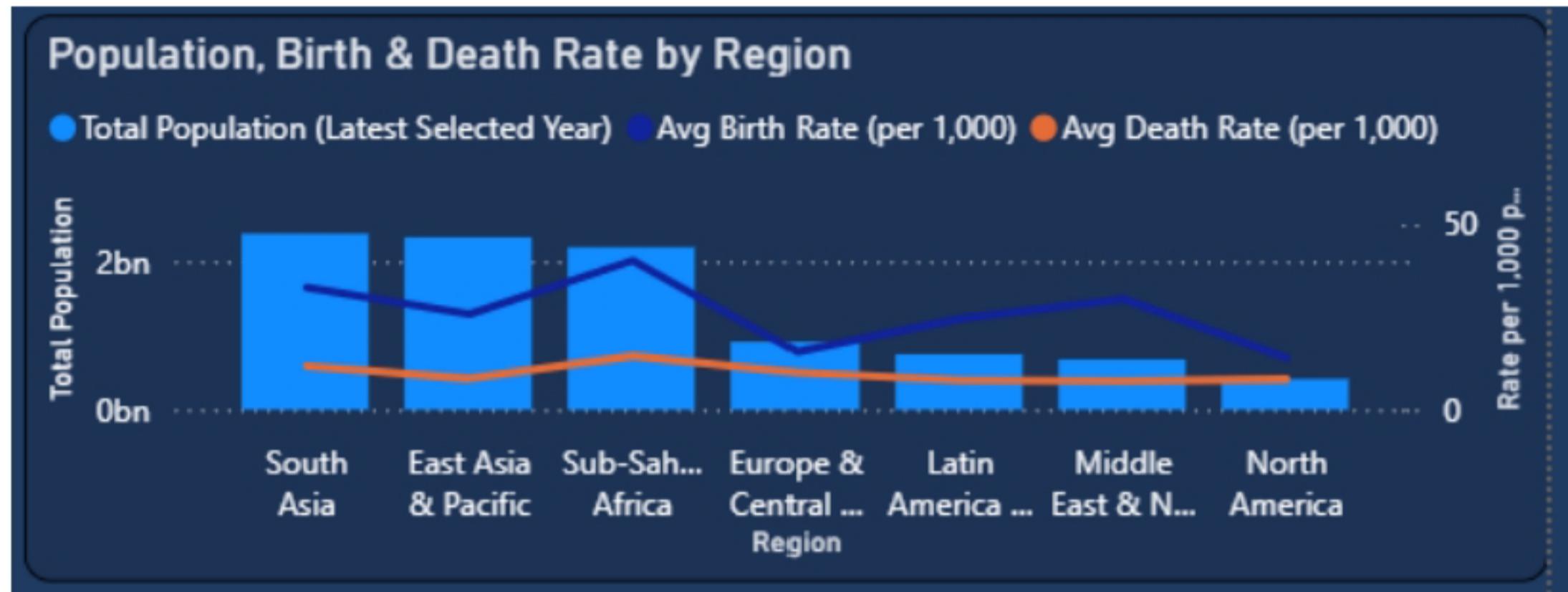


Purpose: quantify global **urbanisation**.

Theory & practice:

- With only **two categories**, a pie chart is efficient; percentage labels are instantly legible.
- Alternative would be a 100% stacked bar; pie chart was chosen for symmetry with Income Group and space economy.

4.4 Combo chart — Region comparison (Population + Birth/Death rates)



Purpose: capture **demographic transition**: population scale alongside fertility/mortality.

Theory:

- Position along a common scale** (bars) gives accurate magnitude comparisons.
- Lines on a secondary axis** encode continuous rates; dual axis is justified by differing units but is clearly titled: left *Total Population*, right *Rate per 1,000 people*.
- Gestalt proximity** and **colour semantics** aid interpretation.

4.5 Line chart — Trend over time



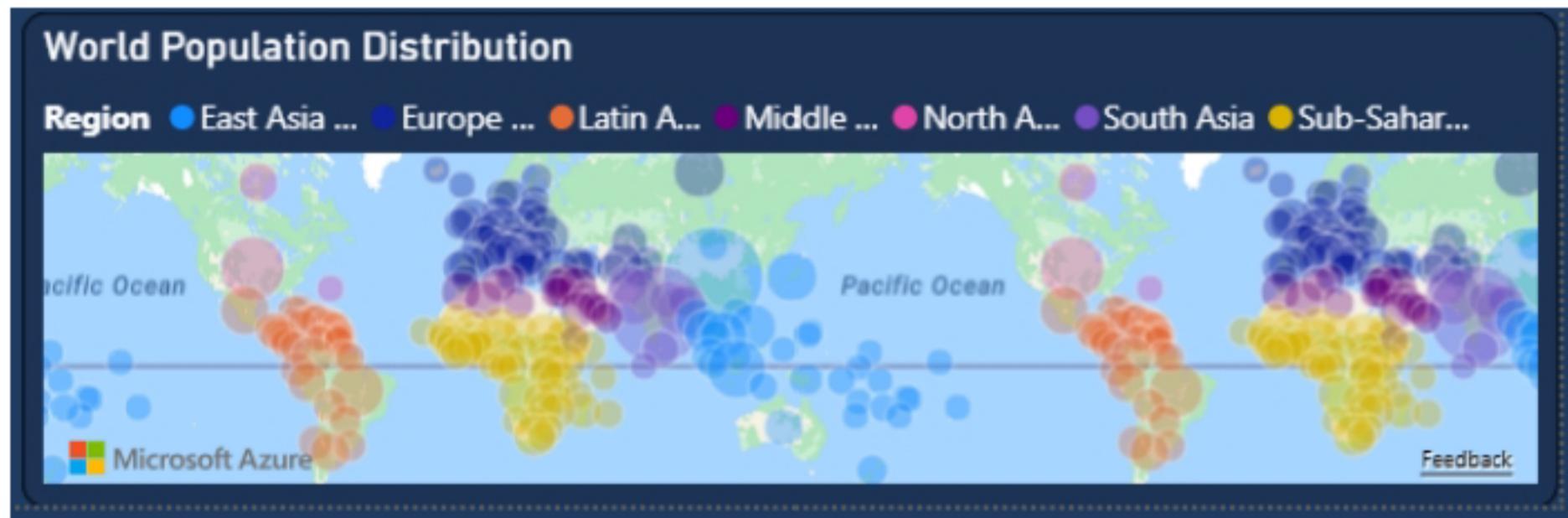
Purpose: long-run **time-series**; supports switching between metrics (Total Population, Female/Male, Growth %, Life Expectancy).

Theory:

- Lines maximise the accuracy of trend perception (position along a continuous scale).

- **Shneiderman's mantra** (overview → filter → details on demand): the metric selector (field parameter) lets users adapt the visual without adding more charts.
 - **Forecast** feature projects near-future values with an explicit confidence band to communicate uncertainty.
-

4.6 Bubble map — World population distribution



Purpose: spatial distribution by country with size encoding **Total Population**.

Theory:

- Maps leverage **spatial schemata** where location meaning is inherent; bubbles avoid choropleth pitfalls when values are not density-based.
 - Tooltips surface extra attributes (Country, Population, Growth %), fulfilling **details on demand**.
-

5. Layout, formatting & composition

5.1 Page structure

The page follows a deliberate **top-down analytic flow**:

1. **Header & filters** (title; Region/Country/Year slicers) → immediate context setting.
2. **KPIs** (cards) → headline measures for the current slice.
3. **Comparisons** (donut + pie chart + combo chart) → categorical insights by income/urbanisation/region.
4. **Trends & geography** (line + map) → time and space perspectives.

This mirrors **Fischer's** and **Shneiderman's** prescriptions: start with overview, allow filtering, and provide details on demand. It also respects **Hichert SUCCESS** ideas: a clear story with consistent notation.

5.2 Dark theme and typography

- **Background # 1F3B61** (very dark navy); **visual panels #1B2B4B**.

- Primary text # FFFFFF, secondary # FFFFFF.
- Sans-serif font (Segoe UI/Open Sans) for legibility; large value sizes in cards; axis labels 10–11pt.

Why dark: it increases contrast and attenuates chrome so the data pops. Accessibility is managed by ensuring **contrast ratio > 4.5** for text on panel backgrounds.

5.3 Interaction & filtering

- Cross-filter direction is **single** (dimension → fact) to avoid ambiguous filter paths.
- Visual interactions tuned so slicers filter all visuals; the metric selector only affects the trend chart.
- A reset state (All Regions, All Countries, full Year range) is maintained for export screenshots.

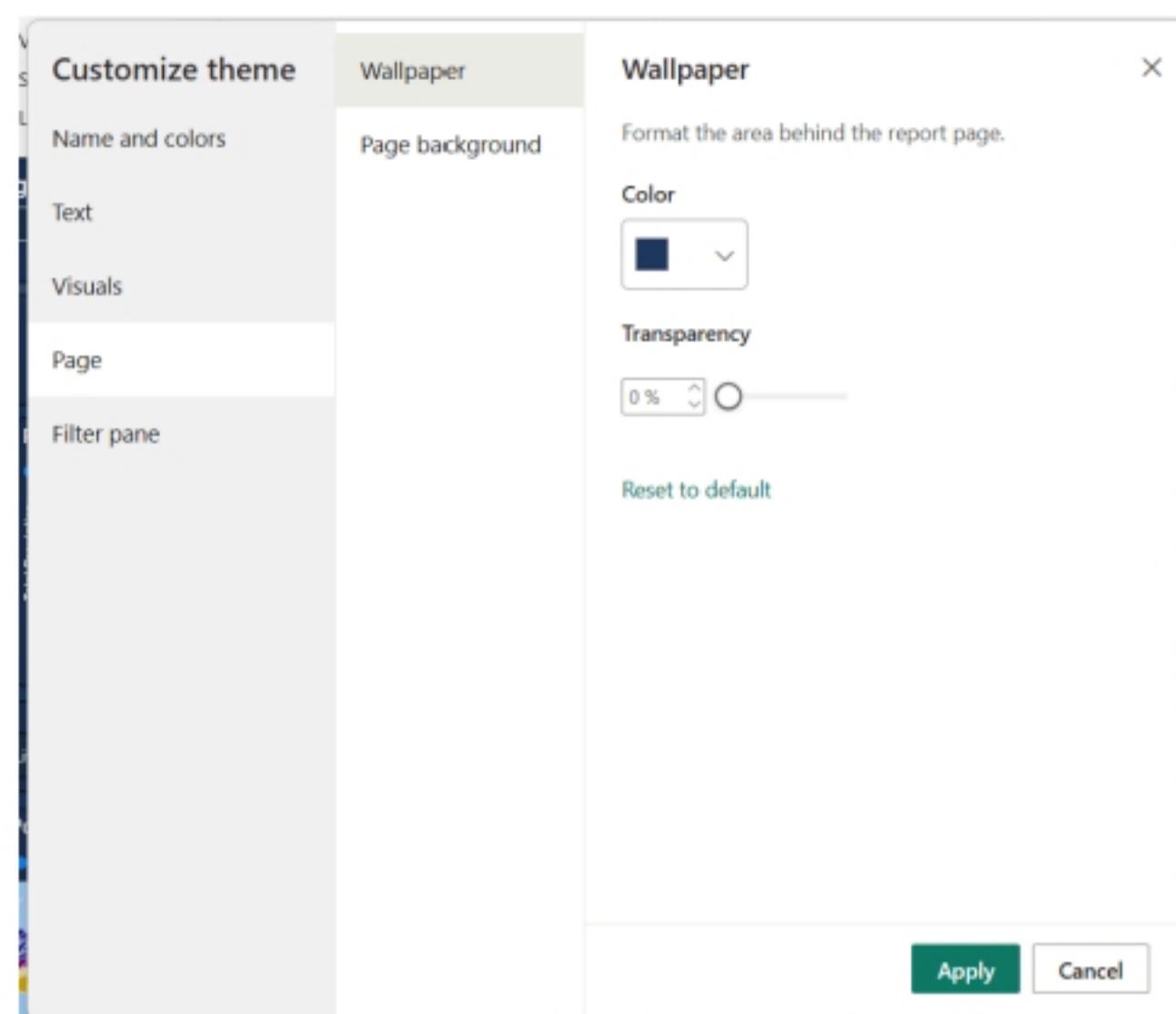
5.4 Performance & maintainability

- Star schema minimises relationships, cutting query ambiguity.
- Measures (not calculated columns) are used for heavy logic, reducing storage and improving reuse.

6. Build Walk-through (step-by-step)

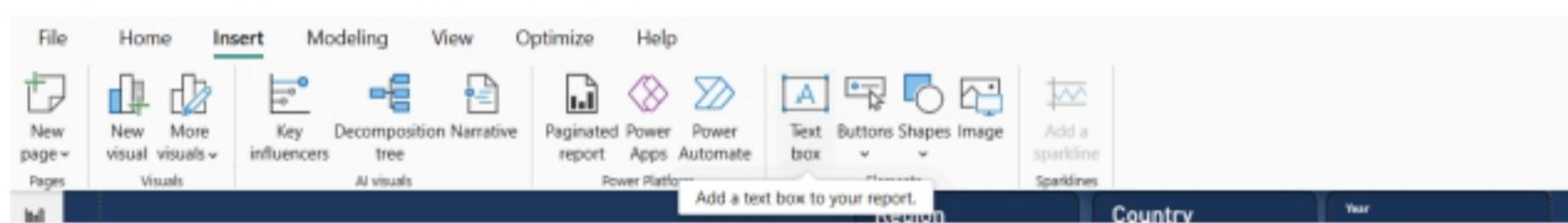
1. Create theme

- Customise colors (*View → Themes → Customize*).

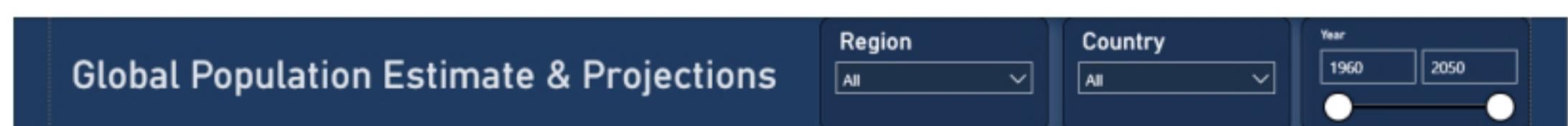
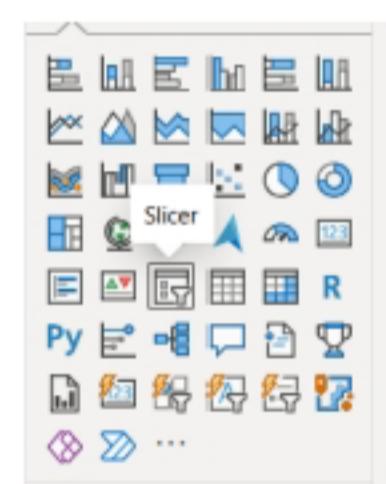


2. Lay out header

- Add a text box for the title band; insert text “Global Population Estimate & Projections”.

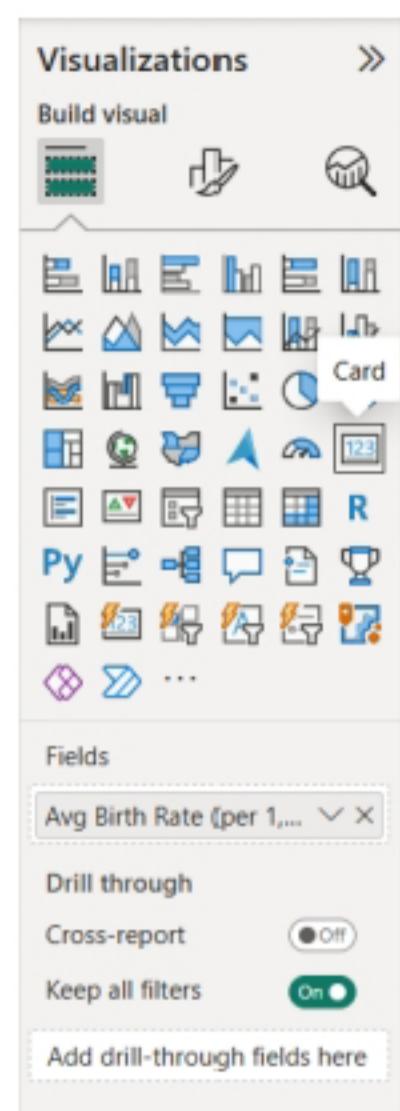


- Add slicers: **Region** (List of economies), **Country**, **Year** (Year table; “Between”/slider).

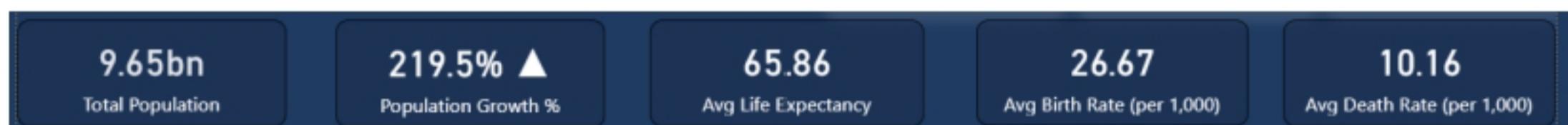


3. Create KPI cards (5 cards)

- Values:** [Total Population], [Population Growth %], [Avg Life Expectancy], [Avg Birth Rate (per 1,000)], [Avg Death Rate (per 1,000)].

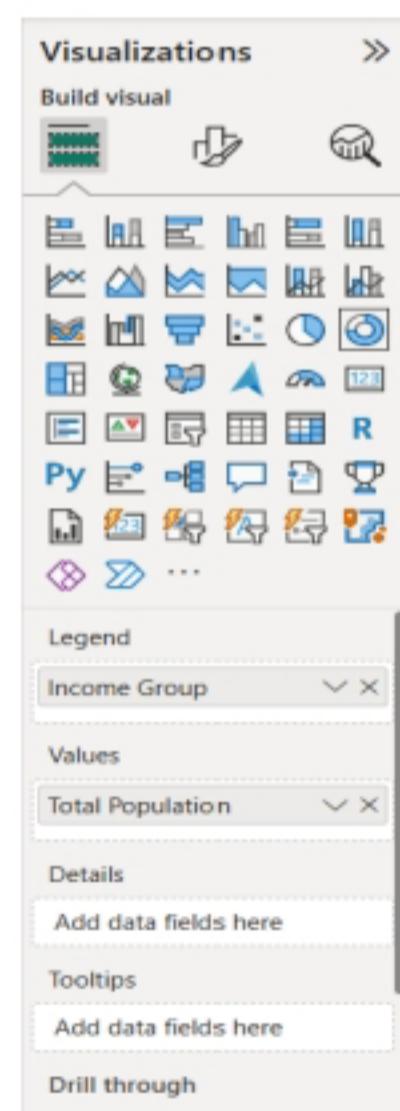


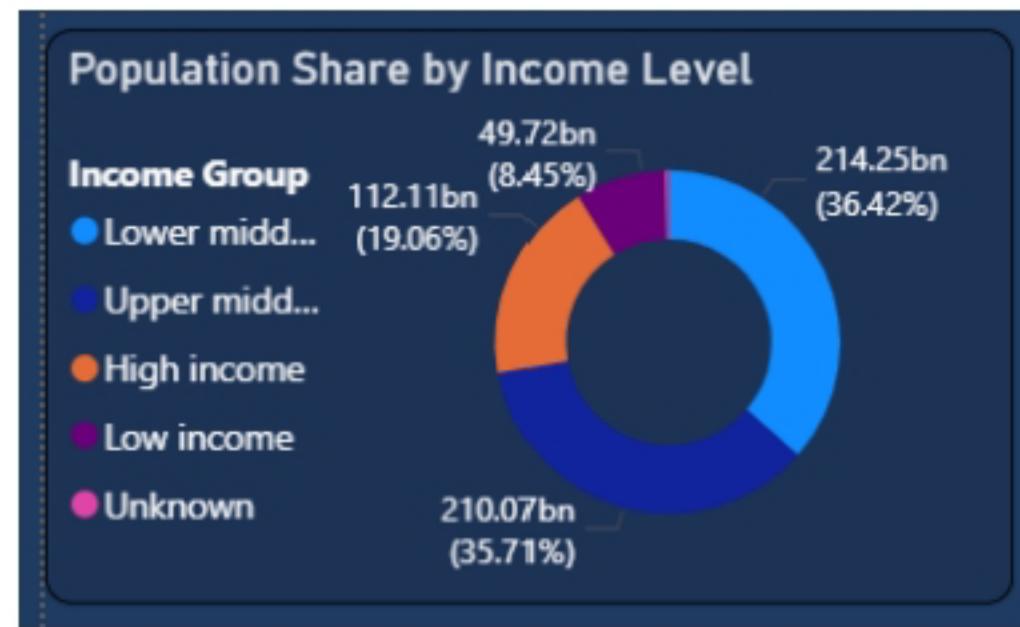
- o Set display units and decimals.



4. Donut — Income Group

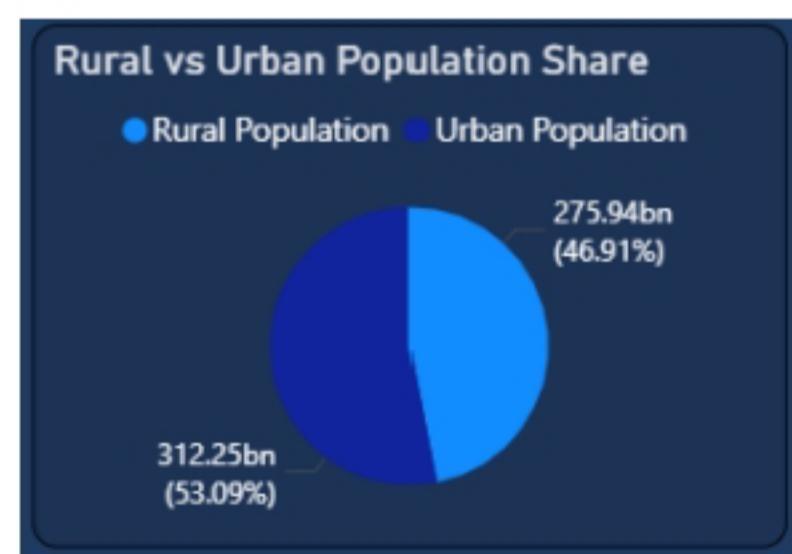
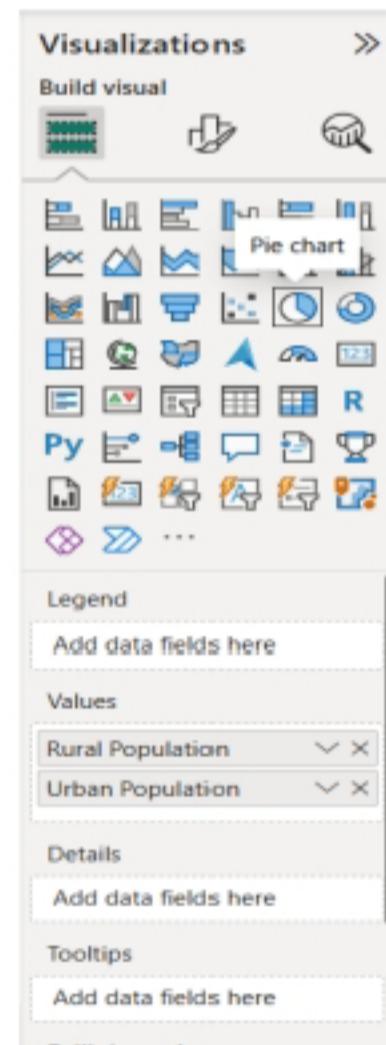
- o Legend = *Income Group*; Value = [Total Population]; show % labels; consistent slice colours.





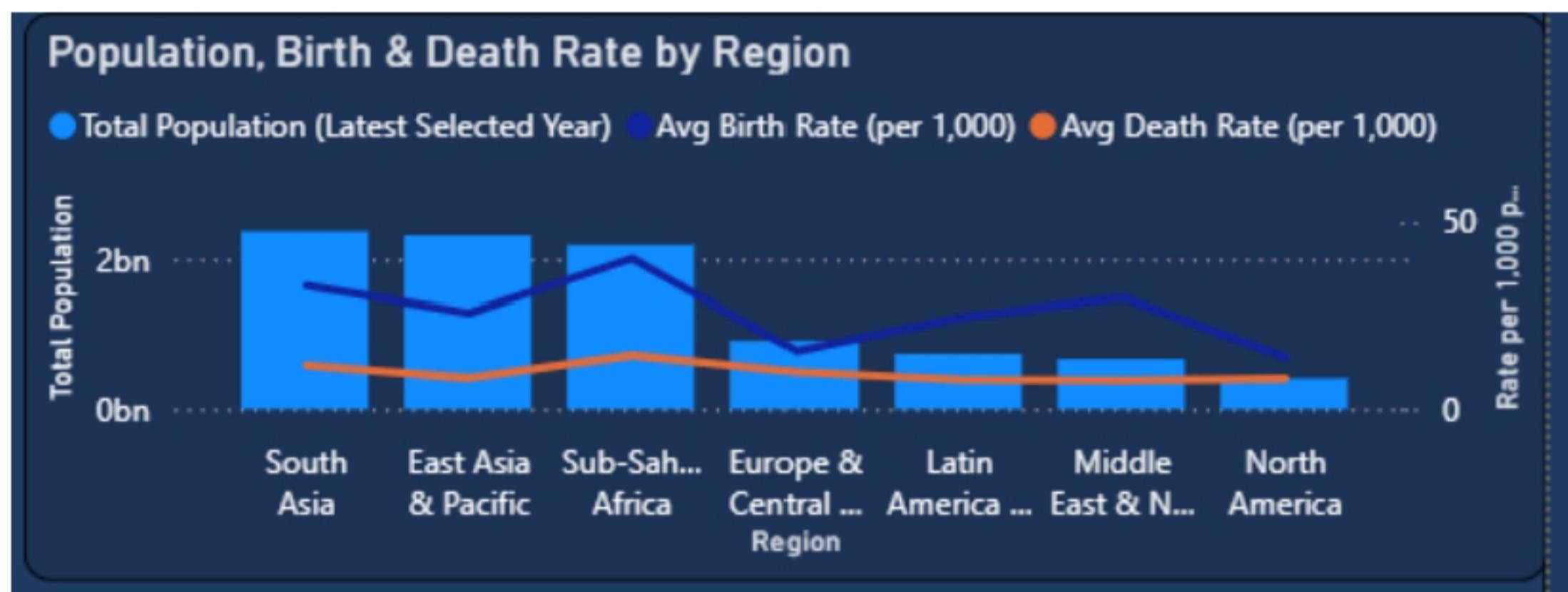
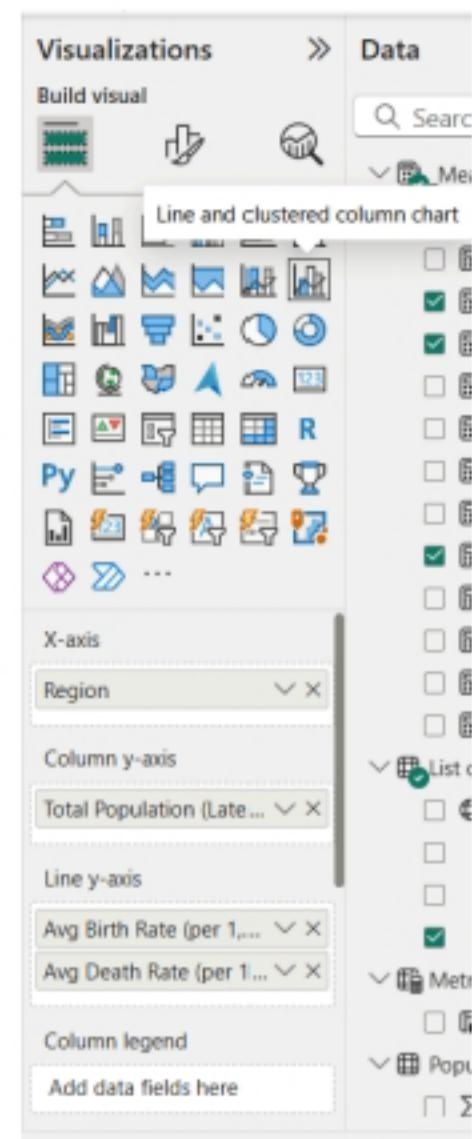
5. Pie chart — Rural vs Urban

- Convert values to two measures and add them to values; show %; keep legend simple.



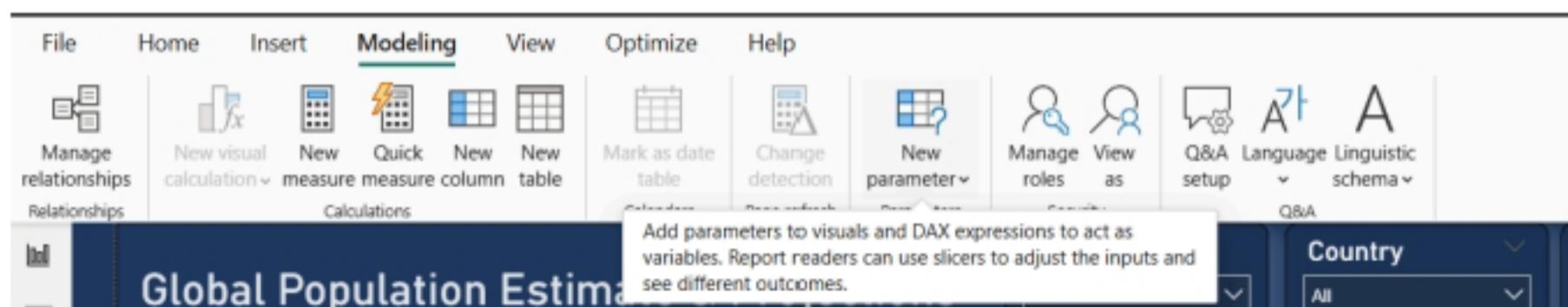
6. Combo chart — Regional comparison

- Axis = *Region*; Column = [Total Population]; Line = [Avg Birth Rate], [Avg Death Rate] (secondary axis).

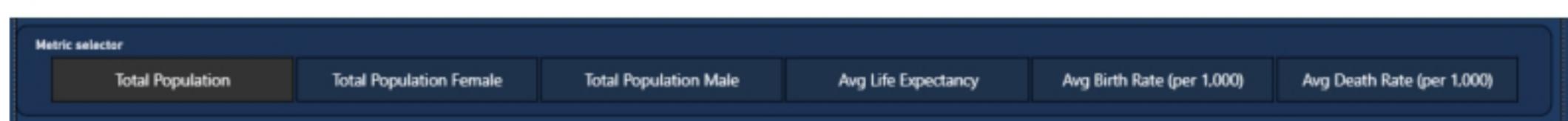


7. Metric selector (field parameter)

- *Modeling → New parameter → Fields; include measures (Total Population, Female, Male, Life Expectancy, Population Growth %, etc.).*

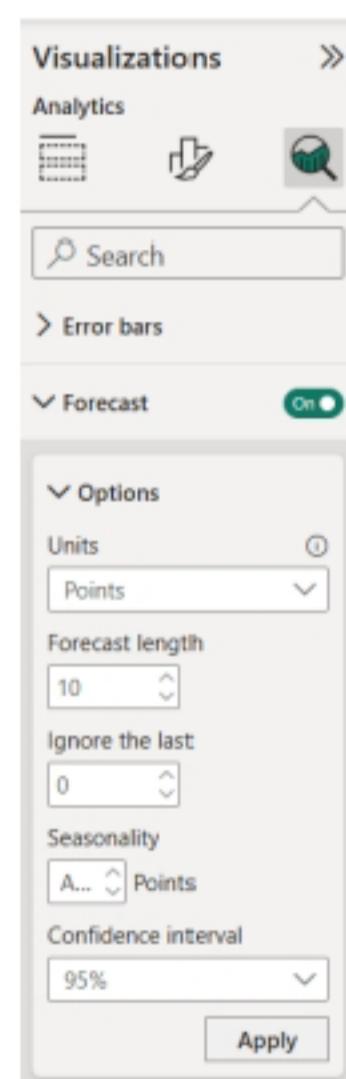
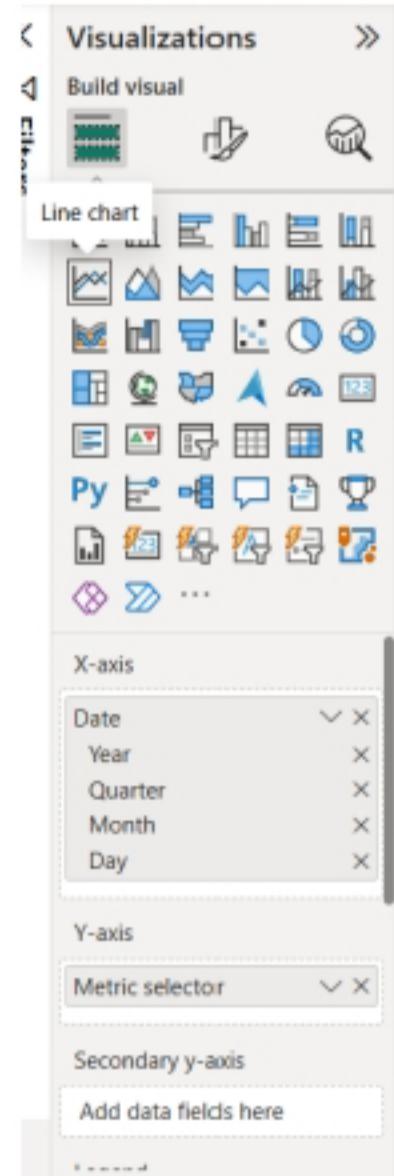


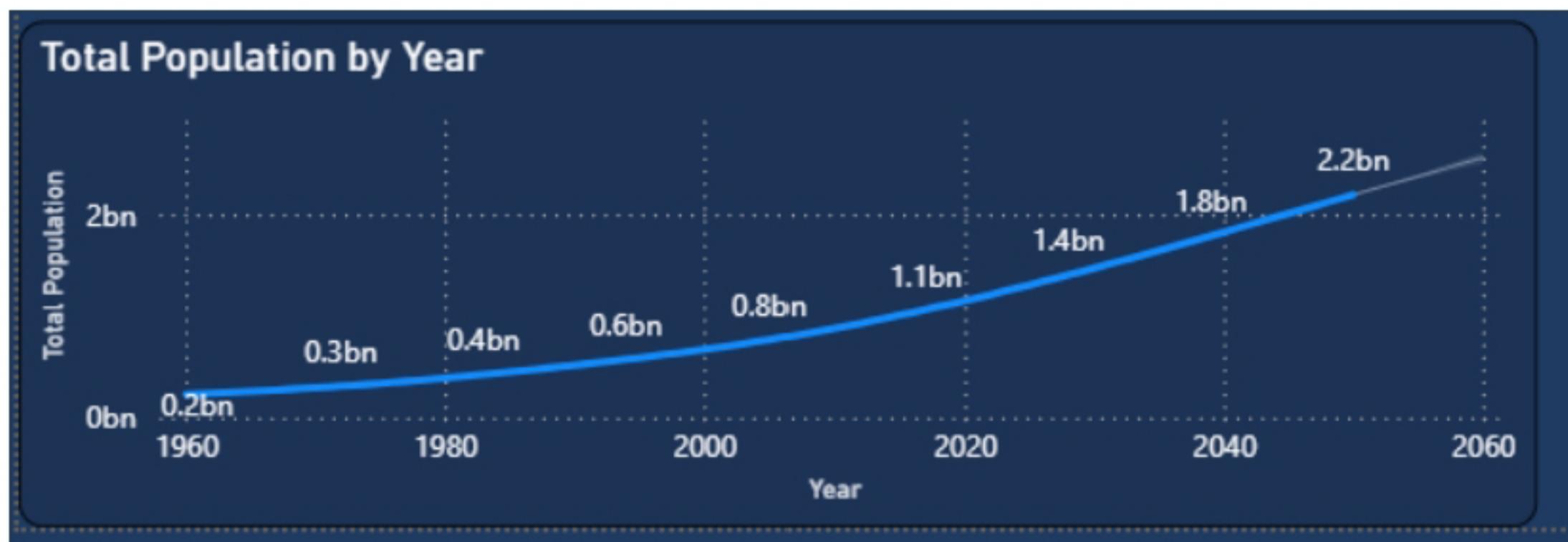
- Place parameter slicer above the line chart.



8. Line chart — Trend

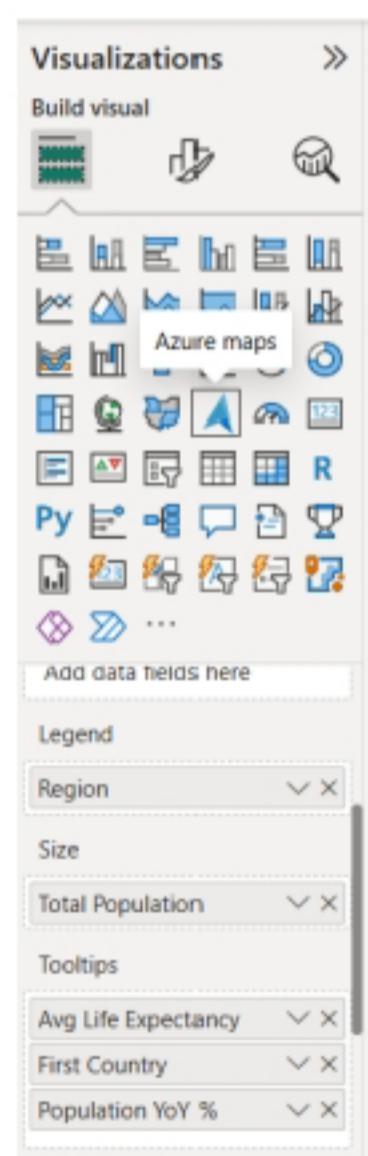
- o X = Year[Date]; Y = **parameter measure**; enable forecast (10 years, 95% CI).
- o Add markers; Y-axis title = measure name (auto from parameter).





9. Map

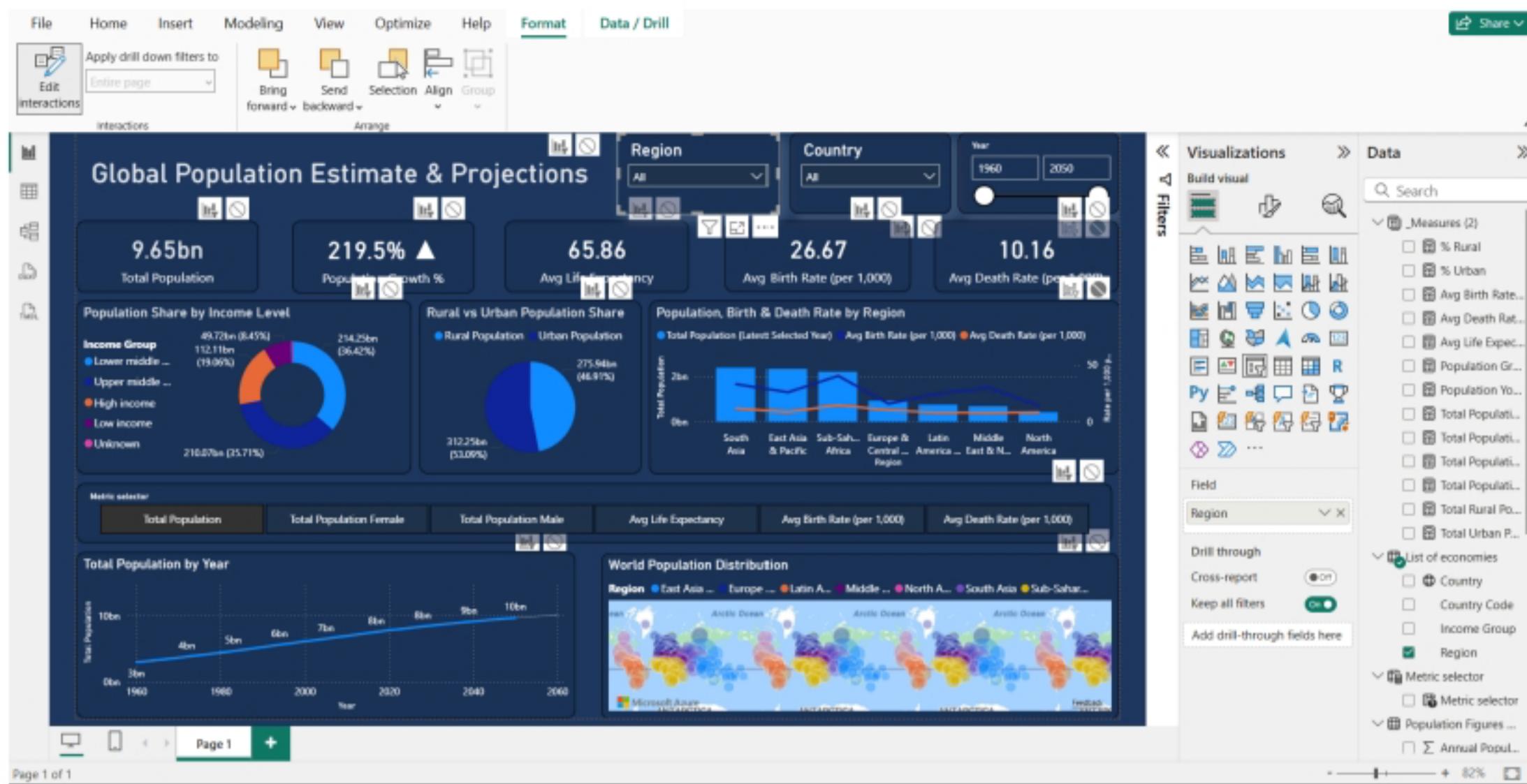
- Location = Country; Bubble size = [Total Population]; Color (optional) = Region; tooltip includes Growth %.
- Bubble transparency 40–50% to avoid overpowering.



10. Interactions

- *Format → Edit interactions*; ensure slicers affect all visuals.

- Validate totals after filtering (e.g., Region=South Asia).



7. Findings & Insights

Using the dashboard:

- **Global trajectory:** total population rises steadily from ~3B in 1960 to ~9–10B by 2050, though growth rate **slows** after ~1990s, consistent with declining fertility.
- **Regional patterns:** **South Asia** and **Sub-Saharan Africa** drive the bulk of future growth; **Europe & Central Asia** is flat or declining in some countries; **East Asia & Pacific** rises but plateaus.
- **Income group composition:** **Upper-middle** and **Lower-middle** income groups account for the largest shares, reflecting China/India and emerging economies; **High income** is smaller but stable.
- **Urbanisation:** global urban share now exceeds **50%** and continues to climb; urban growth is especially marked in middle-income regions.
- **Vital rates:** **Birth rates** decline across regions; **Death rates** fall but can tick upward in aging populations, supporting the **demographic transition** narrative.



8. Critical Evaluation

8.1 What works well

- **Clarity & narrative:** the layout implements an **overview → comparison → trend → geography** storyline that assessors can grasp quickly.
- **Interactivity:** Region/Country/Year slicers and the **metric selector** enable deep exploration without cluttering the page with extra visuals.
- **Theoretical grounding:** visual choices align with **Cleveland & McGill** (perceptual accuracy), **Few** (dashboard clarity), **Tufte** (data-ink), and **Shneiderman** (interaction).
- **Performance:** a tidy star schema and measure-heavy calculations keep the model responsive.
- **Reusability:** measures and display folders are organised; theme centralises visual standards.

8.2 Limitations

- **Data completeness:** early-year gaps in life expectancy and rate indicators may under-represent some countries; nulls are intentionally left un-imputed.
- **Forecast method:** Power BI's built-in forecast is univariate and heuristic; it does not include causal drivers (fertility policies, migration, shocks).
- **Map occlusion:** bubble overlap at global scale can hide small countries (Zoom and tooltips mitigate, but not eliminate).

8.3 Future enhancements

- **Hierarchical drill-down:** Region → Country → (optional) Major city groups if data available; or use a **decomposition tree** for “what changed” analysis.
 - **Tooltip pages:** rich on-hover mini-cards with all KPIs and a tiny sparkline per country.
 - **Small multiples:** regional small-multiple line charts to compare shapes without shared axes complications.
 - **Scenario overlays:** add alternative projection paths (low/medium/high fertility) if the source provides them.
 - **Thematic extensions:** join to GDP per capita, CO₂ emissions, or health expenditure to explore socio-demographic links.
-

9. Conclusion

This project transforms the Global Population Estimates & Projections data into a **concise, interactive, and theory-informed** dashboard. Through careful **data preparation**, a clean **star schema**, and a set of robust **DAX measures**, the report provides a faithful view of how the world’s population has evolved since 1960 and how it is expected to change by 2050. Visual decisions are grounded in well-known principles: **accuracy of encodings** (Cleveland & McGill), **economy of design** (Tufte), **clarity of dashboards** (Few), and **effective interaction** (Shneiderman). The result is a tool that supports both macro-level understanding and targeted country exploration.

While limitations exist—especially around early-period missingness and simple forecasting—the dashboard achieves its objectives: **trend communication, comparative insights, and accessible exploration across region, income group, and time**. The documented build process, advanced features (field parameters, forecasting, conditional formatting), and tidy model design collectively meet the assessment brief’s expectations for technical competency and reflective practice.
