**TREN**

It’s **2025**. You are the leader of a government crew called **T**rade **R**esilience & **E**conomic **N**etworks, A coalition of **25 nations** — ranging from economic superpowers to developing economies — has asked your analytics team for help.  
Over the past two decades, these nations have faced:

* Global recessions
* Trade disputes
* Climate disasters
* Pandemics
* Political instability

They’ve handed you **historical economic, trade, demographic, agricultural, and welfare data** for the years **2000–2024** and asked:

“What will our economies look like in 2030 under different crisis and policy scenarios? And what can we do now to strengthen our resilience and protect our people?”

Your mission: **Clean, integrate, and model** this data to predict and recommend the best economic strategies for each country.

**Countries in the Dataset**

India, USA, Russia, France, Germany, Italy, China, Japan, Argentina, Portugal, Spain, Croatia, Belgium, Australia, Pakistan, Afghanistan, Israel, Iran, Iraq, Bangladesh, Sri Lanka, Canada, UK, Sweden, Saudi Arabia.

**Dataset Overview**

**1. Trade Data (Exports & Imports)**

* Covers **2000–2024** in two time segments (2000–2012 and 2013–2024).
* By value (USD), volume (tonnes), and trade partner.

**2. Core Economic Indicators**

* GDP, GDP per capita, GDP growth, inflation, trade % of GDP, government spending % of GDP, etc.

**3. Crop & Livestock**

* Food production index, yield, livestock counts, agricultural exports/imports.

**4. Disasters**

* Year, type, severity index, estimated damages (USD), lives lost, recovery years.

**5. Employment & Unemployment**

* Employment-to-population ratios, unemployment %, labour force participation.

**6. Population & Demographics**

* Total population, urbanization %, population growth %, median age.

**7. Resilience Metrics**

* Vulnerability score, trade diversification index, disaster recovery score.

**8. Social & Welfare**

* Poverty rates ($2.15/day, $3.65/day, national line), life expectancy, Gini index, HDI.

**Challenge Tasks**

1. **Data Integration & Cleaning**
   * Merge all thematic datasets into a single analysis framework by **Country–Year**.
   * Handle missing data, reconcile conflicting numbers, normalize units.
2. **Feature Engineering**
   * Build composite indexes:
     + **Trade Dependency Index**
     + **Resilience Score**
     + **Spending Efficiency**
     + **Shock Impact Score**
3. **Modeling & Forecasting**
   * Predict GDP growth, poverty rates, and trade resilience for **2030** under:
     + Baseline (no policy change)
     + Increased social spending
     + Trade diversification policies
     + Global crisis scenario (disaster or recession)
4. **Visualization & Insights**
   * Produce heatmaps, trade network graphs, and shock maps for each country.
   * Identify the **3 biggest vulnerabilities** for each nation.
5. **Policy Recommendations**
   * Suggest country-specific strategies for improving resilience by 2030.
   * Simulate “what-if” scenarios to quantify impact.

**Dataset Structure:**

**core\_economic\_indicators.csv**

* **imports\_goods\_services\_gdp\_pct**: The value of all goods and services imported into a country, expressed as a percentage of Gross Domestic Product (GDP).
* **exports\_goods\_services\_gdp\_pct**: The value of all goods and services exported from a country, expressed as a percentage of GDP.
* **trade\_gdp\_pct**: The sum of exports and imports of goods and services, measured as a percentage of GDP.
* **inflation\_consumer\_prices\_pct**: The annual percentage change in the cost to the average consumer of acquiring a fixed basket of goods and services (Consumer Price Index).
* **gdp\_growth\_pct**: The annual percentage growth rate of GDP at market prices, based on constant local currency.
* **gdp\_per\_capita\_current\_usd**: GDP divided by midyear population, in current U.S. dollars.
* **gdp\_current\_usd**: Gross Domestic Product at market prices, in current U.S. dollars.

**crop\_and\_livestock.csv**

* **Domain Code**: A short alphanumeric code representing the statistical domain or category of data (e.g., crops, livestock).
* **Domain**: The descriptive name of the statistical domain (e.g., "Crops and livestock products").
* **Area Code (M49)**: The standardized numeric code assigned to the country or region, based on the United Nations M49 classification.
* **Area**: The name of the country, territory, or region where the data applies.
* **Element Code**: A numeric code representing a specific statistical element (e.g., production, yield, export quantity).
* **Element**: A descriptive label for the statistical element represented by the **Element Code**.
* **Item Code (CPC)**: A numeric code based on the UN Central Product Classification (CPC) system, identifying the product (crop or livestock).
* **Item**: The descriptive name of the product (e.g., "Wheat", "Cattle").
* **Year Code**: A numeric code representing the year of data collection or reporting.
* **Year**: The calendar year for the observation.
* **Unit**: The unit of measurement used for the value (e.g., tonnes, head, kilograms/hectare).
* **Value**: The recorded statistical value corresponding to the specified item, year, and area.
* **Flag**: A symbol or letter providing additional information about the observation, such as data quality or estimation status.
* **Flag Description**: A descriptive note explaining the meaning of the **Flag**.
* **Note**: Any additional remarks or context relevant to the observation, such as special conditions or anomalies in the data.

**disaster.csv**

* **DisNo.:** Unique disaster identifier assigned by the dataset provider; used to track a single event across updates and sources.
* **Historic:** Indicator showing whether the record is a historical entry or outside the primary reporting period (e.g., Yes/No).
* **Classification Key:** Internal key linking the event to its taxonomy (group, subgroup, type, subtype); used for standardized classification.
* **Disaster Group:** Top-level hazard category (e.g., Natural, Technological, Biological).
* **Disaster Subgroup:** Secondary category refining the group (e.g., Hydrological, Meteorological, Geophysical).
* **Disaster Type:** Specific hazard type within the subgroup (e.g., Flood, Tropical cyclone, Earthquake).
* **Disaster Subtype:** Further specification of the disaster type (e.g., Flash flood, Storm surge, Riverine flood).
* **External IDs:** Cross-references to external databases or identifiers (e.g., GLIDE numbers, national IDs).
* **Event Name:** Common or official name of the event as reported (e.g., “Cyclone Idai”).
* **ISO:** Standardized country code for the affected country (typically ISO 3166 alpha-3).
* **Country:** Name of the primary affected country.
* **Subregion:** Geographical subregion designation for the country (e.g., Southern Africa), typically per a standard geoscheme.
* **Region:** Broader geographical region (e.g., Africa, Asia).
* **Location:** Free-text description of the impacted area(s) within the country (e.g., cities, districts, landmarks).
* **Origin:** Source or cause classification (e.g., Natural, Technological, Conflict-related).
* **Associated Types:** Additional or cascading hazard types linked to the event (e.g., Landslide following earthquake).
* **OFDA/BHA Response:** Indicator whether USAID’s humanitarian arm responded or provided assistance (e.g., Yes/No).
* **Appeal:** Indicator whether an international appeal for assistance was launched (e.g., Yes/No).
* **Declaration:** Indicator whether a formal emergency/disaster declaration was made by authorities (e.g., Yes/No).
* **AID Contribution ('000 US$):** Reported aid contributions associated with the event, in thousands of U.S. dollars.
* **Magnitude:** Numeric measure of event strength or intensity (context depends on hazard).
* **Magnitude Scale:** The scale corresponding to Magnitude (e.g., Richter, MMI, Saffir–Simpson, VEI, SPI).
* **Latitude:** Latitude of the event’s centroid or principal impact location (decimal degrees).
* **Longitude:** Longitude of the event’s centroid or principal impact location (decimal degrees).
* **River Basin:** Name of the river basin associated with the event (relevant for floods and hydrological hazards).
* **Start Year:** Calendar year when the event began.
* **Start Month:** Month when the event began (1–12; may be null if unknown).
* **Start Day:** Day of month when the event began (1–31; may be null if unknown).
* **End Year:** Calendar year when the event ended.
* **End Month:** Month when the event ended (1–12; may be null if ongoing/unknown).
* **End Day:** Day of month when the event ended (1–31; may be null if ongoing/unknown).
* **Total Deaths:** Total number of reported fatalities attributed to the event.
* **No. Injured:** Total number of people reported injured.
* **No. Affected:** Number of people requiring immediate assistance during the emergency (e.g., basic survival needs).
* **No. Homeless:** Number of people rendered homeless due to the event.
* **Total Affected:** Aggregate of affected population categories as defined by the provider (typically includes injured, affected, and homeless).
* **Reconstruction Costs ('000 US$):** Estimated reconstruction expenditures, in thousands of current U.S. dollars at time of estimate.
* **Reconstruction Costs, Adjusted ('000 US$):** Reconstruction costs adjusted for inflation to constant U.S. dollars using the provided CPI/deflator.
* **Insured Damage ('000 US$):** Estimated insured losses, in thousands of current U.S. dollars at time of estimate.
* **Insured Damage, Adjusted ('000 US$):** Insured losses adjusted for inflation to constant U.S. dollars using the provided CPI/deflator.
* **Total Damage ('000 US$):** Estimated total economic damages, in thousands of current U.S. dollars at time of estimate.
* **Total Damage, Adjusted ('000 US$):** Total damages adjusted for inflation to constant U.S. dollars using the provided CPI/deflator.
* **CPI:** Consumer Price Index or deflator value used for inflation adjustments for the event’s monetary figures.
* **Admin Units:** List of affected administrative units (e.g., states, provinces, districts).
* **Entry Date:** Date the event record was first created in the database.
* **Last Update:** Date of the most recent modification to the event record.

**resilience.csv**

**current\_account\_balance\_gdp\_pct**: The net flow of goods, services, income, and current transfers as a percentage of Gross Domestic Product (GDP). A positive value indicates a surplus; a negative value indicates a deficit.

**external\_debt\_stocks\_gni\_pct**: The total public and private debt owed to nonresidents, repayable in internationally accepted currencies, goods, or services, expressed as a percentage of Gross National Income (GNI).

**fdi\_net\_inflows\_gdp\_pct**: Foreign direct investment net inflows (new investment inflows minus disinvestment) expressed as a percentage of GDP; measures the attractiveness of the economy to foreign investors.

**social\_and\_welfare.csv**

* urban\_population\_pct\_total – Percentage of the total population living in urban areas, reflecting urbanization trends.
* population\_growth\_annual\_pct – Year-over-year percentage change in total population, showing demographic expansion or contraction.
* unemployment\_total\_pct\_labor\_force – Unemployed persons as a percentage of the total labor force, based on modeled ILO estimates.
* life\_expectancy\_birth\_total\_yrs – Average number of years a newborn is expected to live, given prevailing mortality rates for all ages and sexes combined.
* life\_expectancy\_birth\_female\_yrs – Average number of years a newborn female is expected to live under current mortality rates.
* life\_expectancy\_birth\_male\_yrs – Average number of years a newborn male is expected to live under current mortality rates.
* gini\_index – Measure of income inequality ranging from 0 (perfect equality) to 100 (perfect inequality).
* poverty\_headcount\_ratio\_3ppp\_pct – Percentage of the population living on less than $3.00 per day at 2021 PPP (purchasing power parity).

**EXPORT.CSV**

* **typeCode**:
  + Refers to the type of data being reported.
  + In this dataset, "C" likely stands for "Commodity" or "Trade in Goods."
* **freqCode**:
  + Indicates the frequency of the data.
  + "A" typically means "Annual" data.
* **refPeriodId**:
  + A unique identifier for the reference period.
  + It combines the year and potentially other period details (e.g., "20000101" could refer to January 1, 2000, or simply the start of the annual period for 2000).
* **refYear**:
  + The reference year for the trade data.
* **refMonth**:
  + The reference month for the trade data.
  + "52" might be a special code indicating annual data or a specific period within the year not directly corresponding to a calendar month.
* **period**:
  + The specific period to which the data refers.
  + In this case, it's the year "2000".
* **reporterCode**:
  + A numerical code identifying the reporting country or economy.
* **reporterISO**:
  + The ISO 3166-1 alpha-3 country code for the reporting country.
  + Examples: "ALB" for Albania, "DZA" for Algeria, "AND" for Andorra.
* **reporterDesc**: A descriptive name of the reporting country or economy.
* **flowCode**:
  + A code indicating the type of trade flow.
  + "X" typically stands for "Export".
* **flowDesc**:
  + A descriptive name of the trade flow.
  + "Export" indicates goods leaving the reporting country.
* **partnerCode**:
  + A numerical code identifying the partner country or economy in the trade.
* **partnerISO**:
  + The ISO 3166-1 alpha-3 country code for the partner country.
  + Examples: "AFG" for Afghanistan, "AUS" for Australia, "BEL" for Belgium.
* **partnerDesc**:
  + A descriptive name of the partner country or economy.
* **partner2Code**:
  + A secondary partner code.
  + "0" and "W00" often indicate "World" or "Total" when the trade is with all countries, not a specific one.
* **partner2ISO**:
  + The ISO 3166-1 alpha-3 country code for the secondary partner.
  + "W00" typically represents "World".
* **partner2Desc**:
  + A descriptive name of the secondary partner.
  + "World" indicates trade with all countries globally.
* **classificationCode**:
  + The code for the classification system used for commodities.
  + "H1" or "H0" likely refers to different versions of the Harmonized System (HS) of commodity classification.
* **classificationSearchCode**:
  + A code used for searching within the classification system.
  + "HS" confirms the Harmonized System.
* **isOriginalClassification**:
  + A boolean value indicating whether the classification is original.
  + "true" means it is the original classification.
* **cmdCode**:
  + The commodity code.
  + "TOTAL" indicates that the data represents "All Commodities" or the total trade value across all commodities.
* **cmdDesc**:
  + A descriptive name of the commodity.
  + "All Commodities" means the data aggregates all goods traded.
* **aggrLevel**:
  + The aggregation level of the commodity data.
  + "0" likely means the highest level of aggregation (e.g., total commodities).
* **isLeaf**:
  + A boolean value indicating if the commodity is a leaf node in the classification hierarchy.
  + "false" for "TOTAL" means it's not a specific, granular commodity.
* **customsCode**:
  + A code related to customs procedures or classifications.
  + "C00" might be a general customs code.
* **customsDesc**:
  + A description of the customs code.
  + "TOTAL CPC" might refer to a total customs product classification.
* **mosCode**:
  + Mode of transport code.
  + "0" might indicate "Total" or "Not specified".
* **motCode**:
  + Mode of transport code.
  + "0" might indicate "Total" or "Not specified".
* **motDesc**:
  + Description of the mode of transport.
  + "TOTAL MOT" likely means "Total Mode of Transport" or all modes combined.
* **qtyUnitCode**:
  + Code for the unit of quantity.
  + "-1" or "N/A" indicates that quantity is not applicable or not provided for this aggregated data.
* **qtyUnitAbbr**:
  + Abbreviation for the unit of quantity.
  + "N/A" means "Not Applicable".
* **qty**:
  + The quantity of goods traded.
  + Empty or "0" values suggest that quantity data is not available or not relevant for these aggregated "TOTAL" commodity entries.
* **isQtyEstimated**:
  + A boolean value indicating if the quantity is estimated.
  + "false" or "true" depending on whether the quantity (if present) was estimated.
* **altQtyUnitCode**:
  + Code for an alternative unit of quantity.
  + "-1" or "N/A" indicates not applicable.
* **altQtyUnitAbbr**:
  + Abbreviation for the alternative unit of quantity.
  + "N/A" means "Not Applicable".
* **altQty**:
  + The alternative quantity of goods traded.
  + Empty or "0" values suggest not applicable.
* **isAltQtyEstimated**:
  + A boolean value indicating if the alternative quantity is estimated.
  + "false" or "true" depending on whether the alternative quantity (if present) was estimated.
* **netWgt**:
  + The net weight of the goods.
  + Empty or "0" values suggest not applicable for these aggregated entries.
* **isNetWgtEstimated**:
  + A boolean value indicating if the net weight is estimated.
  + "false" or "true" depending on whether the net weight (if present) was estimated.
* **grossWgt**:
  + The gross weight of the goods.
  + Empty or "0" values suggest not applicable for these aggregated entries.
* **isGrossWgtEstimated**:
  + A boolean value indicating if the gross weight is estimated.
  + "false" or "true" depending on whether the gross weight (if present) was estimated.
* **cifvalue**:
  + The Cost, Insurance, and Freight (CIF) value of the goods. This is typically used for import values.
  + For exports, it might be empty or the same as FOB value if not specifically tracked.
* **fobvalue**:
  + The Free On Board (FOB) value of the goods. This is typically used for export values, representing the value of goods at the point of shipment.
* **primaryValue**:
  + The primary trade value, which is usually the FOB value for exports and CIF value for imports.
  + In this export-focused dataset, it often mirrors the fobvalue.
* **legacyEstimationFlag**:
  + A flag related to legacy estimation methods.
  + "0" or "4" might indicate different types or levels of estimation.
* **isReported**:
  + A boolean value indicating whether the data was directly reported.
  + "false" or "true" depending on the reporting status.
* **isAggregate**:
  + A boolean value indicating whether the data is an aggregate (summed up) value.
  + "false" or "true" depending on whether the data is aggregated.

**IMPORT.CSV**

**freqCode**:

* + Indicates the frequency of the data.
  + "A" typically means "Annual" data.
* **refPeriodId**:
  + A unique identifier for the reference period.
  + It combines the year and potentially other period details (e.g., "20000101" could refer to January 1, 2000, or simply the start of the annual period for 2000).
* **refYear**:
  + The reference year for the trade data.
* **refMonth**:
  + The reference month for the trade data.
  + "52" might be a special code indicating annual data or a specific period within the year not directly corresponding to a calendar month.
* **period**:
  + The specific period to which the data refers.
  + In this case, it's the year "2000".
* **reporterCode**:
  + A numerical code identifying the reporting country or economy.
* **reporterISO**:
  + The ISO 3166-1 alpha-3 country code for the reporting country.
  + Examples: "ALB" for Albania, "DZA" for Algeria, "AND" for Andorra.
* **reporterDesc**:
  + A descriptive name of the reporting country or economy.
* **flowCode**:
  + A code indicating the type of trade flow.
  + "X" typically stands for "Export".
* **flowDesc**:
  + A descriptive name of the trade flow.
  + "Export" indicates goods leaving the reporting country.
* **partnerCode**:
  + A numerical code identifying the partner country or economy in the trade.
* **partnerISO**:
  + The ISO 3166-1 alpha-3 country code for the partner country.
  + Examples: "AFG" for Afghanistan, "AUS" for Australia, "BEL" for Belgium.
* **partnerDesc**:
  + A descriptive name of the partner country or economy.
* **partner2Code**:
  + A secondary partner code.
  + "0" and "W00" often indicate "World" or "Total" when the trade is with all countries, not a specific one.
* **partner2ISO**:
  + The ISO 3166-1 alpha-3 country code for the secondary partner.
  + "W00" typically represents "World".
* **partner2Desc**:
  + A descriptive name of the secondary partner.
  + "World" indicates trade with all countries globally.
* **classificationCode**:
  + The code for the classification system used for commodities.
  + "H1" or "H0" likely refers to different versions of the Harmonized System (HS) of commodity classification.
* **classificationSearchCode**:
  + A code used for searching within the classification system.
  + "HS" confirms the Harmonized System.
* **isOriginalClassification**:
  + A boolean value indicating whether the classification is original.
  + "true" means it is the original classification.
* **cmdCode**:
  + The commodity code.
  + "TOTAL" indicates that the data represents "All Commodities" or the total trade value across all commodities.
* **cmdDesc**:
  + A descriptive name of the commodity.
  + "All Commodities" means the data aggregates all goods traded.
* **aggrLevel**:
  + The aggregation level of the commodity data.
  + "0" likely means the highest level of aggregation (e.g., total commodities).
* **isLeaf**:
  + A boolean value indicating if the commodity is a leaf node in the classification hierarchy.
  + "false" for "TOTAL" means it's not a specific, granular commodity.
* **customsCode**:
  + A code related to customs procedures or classifications.
  + "C00" might be a general customs code.
* **customsDesc**:
  + A description of the customs code.
  + "TOTAL CPC" might refer to a total customs product classification.
* **mosCode**:
  + Mode of transport code.
  + "0" might indicate "Total" or "Not specified".
* **motCode**:
  + Mode of transport code.
  + "0" might indicate "Total" or "Not specified".
* **motDesc**:
  + Description of the mode of transport.
  + "TOTAL MOT" likely means "Total Mode of Transport" or all modes combined.
* **qtyUnitCode**:
  + Code for the unit of quantity.
  + "-1" or "N/A" indicates that quantity is not applicable or not provided for this aggregated data.
* **qtyUnitAbbr**:
  + Abbreviation for the unit of quantity.
  + "N/A" means "Not Applicable".
* **qty**:
  + The quantity of goods traded.
  + Empty or "0" values suggest that quantity data is not available or not relevant for these aggregated "TOTAL" commodity entries.
* **isQtyEstimated**:
  + A boolean value indicating if the quantity is estimated.
  + "false" or "true" depending on whether the quantity (if present) was estimated.
* **altQtyUnitCode**:
  + Code for an alternative unit of quantity.
  + "-1" or "N/A" indicates not applicable.
* **altQtyUnitAbbr**:
  + Abbreviation for the alternative unit of quantity.
  + "N/A" means "Not Applicable".
* **altQty**:
  + The alternative quantity of goods traded.
  + Empty or "0" values suggest not applicable.
* **isAltQtyEstimated**:
  + A boolean value indicating if the alternative quantity is estimated.
  + "false" or "true" depending on whether the alternative quantity (if present) was estimated.
* **netWgt**:
  + The net weight of the goods.
  + Empty or "0" values suggest not applicable for these aggregated entries.
* **isNetWgtEstimated**:
  + A boolean value indicating if the net weight is estimated.
  + "false" or "true" depending on whether the net weight (if present) was estimated.
* **grossWgt**:
  + The gross weight of the goods.
  + Empty or "0" values suggest not applicable for these aggregated entries.
* **isGrossWgtEstimated**:
  + A boolean value indicating if the gross weight is estimated.
  + "false" or "true" depending on whether the gross weight (if present) was estimated.
* **cifvalue**:
  + The Cost, Insurance, and Freight (CIF) value of the goods. This is typically used for import values.
  + For exports, it might be empty or the same as FOB value if not specifically tracked.
* **fobvalue**:
  + The Free On Board (FOB) value of the goods. This is typically used for export values, representing the value of goods at the point of shipment.
* **primaryValue**:
  + The primary trade value, which is usually the FOB value for exports and CIF value for imports.
  + In this export-focused dataset, it often mirrors the fobvalue.
* **legacyEstimationFlag**:
  + A flag related to legacy estimation methods.
  + "0" or "4" might indicate different types or levels of estimation.
* **isReported**:
  + A boolean value indicating whether the data was directly reported.
  + "false" or "true" depending on the reporting status.
* **isAggregate**:
  + A boolean value indicating whether the data is an aggregate (summed up) value.
  + "false" or "true" depending on whether the data is aggregated.

**Employment\_Unemployment.csv**

* **Country Name**: This column contains the name of the country for which the employment and unemployment data is provided.
* **Country Code**: This column contains the unique three-letter code representing each country.
* **Series Name**: This column describes the specific employment or unemployment indicator being measured. Examples include:
  + "Unemployment with advanced education, male (% of male labor force with advanced education)"
  + "Employment to population ratio, ages 15-24, female (%) (national estimate)"
  + "Unemployment, total (% of total labor force) (modeled ILO estimate)"
* **Series Code**: This column contains a unique code for each series name, often indicating the type of indicator (e.g., SL.UEM for unemployment, SL.EMP for employment to population ratio).
* **2000 [YR2000] - 2024 [YR2024]**: These columns represent the values for the respective employment/unemployment indicators for each year from 2000 to 2024. ".." indicates that data is not available for that specific year and series.

**POPULATION AND DEMOGRAPHICS.CSV**

* **Domain Code**: This column likely contains a short code representing the domain of the data, which is "Annual population" in this file.
* **Domain**: This column specifies the broader category or theme of the data, which is "Annual population".
* **Area Code (M49)**: This column holds the M49 standard code for the geographical area or country.
* **Area**: This column provides the name of the geographical area or country.
* **Element Code**: This column contains a code representing the specific element or type of population data being measured (e.g., Total Population - Both sexes, Total Population - Male, Rural population).
* **Element**: This column describes the specific population metric being reported (e.g., "Total Population - Both sexes", "Total Population - Male", "Total Population - Female", "Rural population", "Urban population").
* **Item Code**: This column contains a code for the item being measured, which is consistently "3010" in this dataset, likely referring to "Population - Est. & Proj.".
* **Item**: This column describes the specific item or type of population data, which is "Population - Est. & Proj." (Population - Estimates & Projections).
* **Year Code**: This column contains a numerical code for the year of the data.
* **Year**: This column indicates the specific year for which the population data is recorded.
* **Unit**: This column specifies the unit of measurement for the 'Value' column, which is "1000 No" (thousands of numbers/people).
* **Value**: This column contains the numerical value of the population for the given element, area, and year, expressed in thousands.
* **Flag**: This column contains a flag indicating certain characteristics of the data, with "X" being the common value in this file.
* **Flag Description**: This column provides a description for the 'Flag' value, which is consistently "Figure from international organizations" in this file.
* **Note**: This column is for any additional notes or comments related to the data. In the provided context, it appears to be empty.

***Problems to Solve***

1. Which countries are most at risk from a single-partner trade collapse?  
    Compute a **Trade Dependency Index**, identify the top 3 vulnerable nations, and simulate GDP impact if their top trade partner’s imports drop by 40% in 2026.
2. Model the cascading trade effects if **China’s exports drop by 25%** in 2028.  
    Identify which 5 countries suffer the greatest GDP percentage loss.

1. Using crop yield + disaster severity data, model how **3 consecutive drought years** would impact agricultural exports for at least 10 countries by 2030.

1. Identify countries most dependent on agricultural imports from only 2–3 partners.  
   Model food security risk if those partners impose export bans.
2. Predict which countries will have **youth unemployment >25%** by 2030 under a global slowdown scenario.

1. Which countries’ export sectors are most at risk from **labour shortages** due to ageing demographics?  
   Simulate productivity impacts if the median age rises by 5,10,15 years.
2. Build a **global trade network graph** for the 25 countries.  
   Identify the most central countries and simulate network disruption if one is removed.
3. Which trade relationships (country-country pairs) have the **highest mutual benefit**?  
   Simulate the GDP impact for both partners if that trade route collapses.

1. Using partner trade concentration metrics, recommend **3 new trade partners** for each country to increase resilience.

1. Predict GDP, poverty, and unemployment in 2030 for each country if:
   1. A disaster hits in 2026 (severity index > 8)
   2. A trade war starts in 2027 cutting global trade volumes by 20%.
2. For each country, model 2030 GDP and poverty rates under:
   1. Best case (trade diversification + resilience investments)
   2. Worst case (recurring disasters + trade concentration).

1. Based on modelled resilience scores, identify which **5 countries are most likely** to rank in the top resilience tier by 2030 — and what key factors drive their success.
2. Formulate an **optimization problem**:
   1. Goal: Minimize the maximum GDP loss for all countries under any single-point failure.
   2. Constraints: Limited budget for new trade deals or infrastructure routes, geopolitical constraints (some countries cannot trade directly), transport distance limits, capacity constraints.
   3. Output: Optimal set of **new trade and infrastructure links** to build.

**Deliverables & Scoring**

| **Category** | **Point** |
| --- | --- |
| Data Integration & Cleaning | 20 |
| Feature Engineering | 20 |
| Modeling Accuracy & Robustness | 25 |
| Insights & Policy Recommendations | 20 |
| Presentation & Visualization | 15 |
| **Total:** 100 points |  |