

BCS Level 4 Diploma Digital Modular Programme for Data Analysts: Synoptic Project Scenarios

Example scenario

Scenario A – Large UK retailer looking to expand into other countries

You are working as a retail Data Analyst and your company has asked you to model the trends in other countries that could help decide where to expand their presence. The retailer sells high-end fashion apparel to high-net worth individuals and focuses on being an ethical trader. Sustainability is a priority. Customers would tend to be well-educated and expect to use online tools to aid their purchases.

The company would expect the chosen countries to have a sustainable growth in both population and income with justification for any forecasting methods used.

The solution should:

- Include the most appropriate countries to expand a high-cost product range into, clearly detailing the criteria of the choices.
- Forecast growth in the chosen countries for the next 10 years.
- Enable the data to be visualised easily by non-technical audiences and senior decision makers.
- The model should be reusable so that it can be use with other datasets.

You will need to produce:

- A clear requirements statement based on stakeholder discussions. (NOTE: Your tutor may fulfil the role of a stakeholder to enable this)
- Details of the processes used for migration and collation of data
- A design and documentation for an appropriate statistical model using a tool of your choice.
- A summary and visualisation of the model output.
- A stakeholder presentation of the results and learnings.

You should use the following datasets for the basis of your analysis -

1. **World Bank Population Estimates and Projections.** A CSV file containing population and demographic estimates and projections from 1960 to 2050. It includes 215 indicators for various population metrics (see Appendix A)
2. **World Bank Environment, Social and Governance Data.** A CSV file containing 71 indicators of various environmental, social and governance categories (see Appendix B)

You can also source other datasets to further your analysis and detail any assumptions you make.

You should produce a presentation that includes the outputs of your analysis and recommendations that could then be delivered to the senior decision makers. As this is the first task that you have been asked to undertake for the organisation you are keen to also use this as an opportunity to evaluate your own skills and identify areas for personal improvement.

Assessment notes – expectation that learners would use the relevant indicators from the ESG data to match the criteria of the retailer and as a stretching potential could include ethical arguments (K15). For planning growth in countries could require some reverse engineering in the provided forecast numbers but also linking the data to other indicators. A good opportunity to use a number of different techniques but learners are clearly directed towards time-series.

Additional suggested scenarios

Scenario B – Transportation company looking to invest in new hubs/routes

Using the global data, analyse key country information that implies economic growth and domestic demand. Added complexity available in looking into links between different countries and optimising the location of the transport hubs. Opportunity to forecast based on time-series data but also could allow use of clustering techniques to model key demand areas.

Scenario C – Engineering company looking to locate new HQ and research facilities

Use of educational and population data with investigation into skilled resource availability.

Scenario D - Medical (overweight, mortality, undernourishment, life expectancy, hospital beds)

Forecasting of growth in medical requirements and population needs for specific countries.

Appendix A – Population Estimates and Projections indicators

Archive file contains six csv files

Population-EstimatesCountry.csv
Population-EstimatesCountry-Series.csv
Population-EstimatesData.csv
Population-EstimatesFootnote.csv
Population-EstimatesSeries.csv
Population-EstimatesSeries-Time.csv

List of indicators in the file Population-EstimatesData.csv

Age dependency ratio (% of working-age population)
Age dependency ratio, old
Age dependency ratio, young
Age population, age 00, female, interpolated
Age population, age 00, male, interpolated
Age population, age 01, female, interpolated
Age population, age 01, male, interpolated
Age population, age 02, female, interpolated
Age population, age 02, male, interpolated
Age population, age 03, female, interpolated
Age population, age 03, male, interpolated
Age population, age 04, female, interpolated
Age population, age 04, male, interpolated
Age population, age 05, female, interpolated
Age population, age 05, male, interpolated
Age population, age 06, female, interpolated
Age population, age 06, male, interpolated
Age population, age 07, female, interpolated
Age population, age 07, male, interpolated
Age population, age 08, female, interpolated
Age population, age 08, male, interpolated
Age population, age 09, female, interpolated
Age population, age 09, male, interpolated
Age population, age 10, female, interpolated
Age population, age 10, male, interpolated
Age population, age 11, female, interpolated
Age population, age 11, male, interpolated
Age population, age 12, female, interpolated
Age population, age 12, male, interpolated
Age population, age 13, female, interpolated
Age population, age 13, male, interpolated
Age population, age 14, female, interpolated

Age population, age 14, male, interpolated
Age population, age 15, female, interpolated
Age population, age 15, male, interpolated
Age population, age 16, female, interpolated
Age population, age 16, male, interpolated
Age population, age 17, female, interpolated
Age population, age 17, male, interpolated
Age population, age 18, female, interpolated
Age population, age 18, male, interpolated
Age population, age 19, female, interpolated
Age population, age 19, male, interpolated
Age population, age 20, female, interpolated
Age population, age 20, male, interpolated
Age population, age 21, female, interpolated
Age population, age 21, male, interpolated
Age population, age 22, female, interpolated
Age population, age 22, male, interpolated
Age population, age 23, female, interpolated
Age population, age 23, male, interpolated
Age population, age 24, female, interpolated
Age population, age 24, male, interpolated
Age population, age 25, female, interpolated
Age population, age 25, male, interpolated
Birth rate, crude (per 1,000 people)
Death rate, crude (per 1,000 people)
Fertility rate, total (births per woman)
Life expectancy at birth, female (years)
Life expectancy at birth, male (years)
Life expectancy at birth, total (years)
Mortality rate, adult, female (per 1,000 female adults)
Mortality rate, adult, male (per 1,000 male adults)
Mortality rate, infant (per 1,000 live births)
Mortality rate, infant, female (per 1,000 live births)
Mortality rate, infant, male (per 1,000 live births)
Mortality rate, neonatal (per 1,000 live births)
Mortality rate, under-5 (per 1,000)
Mortality rate, under-5, female (per 1,000)
Mortality rate, under-5, male (per 1,000)
Net migration
Number of deaths ages 10-14 years
Number of deaths ages 15-19 years
Number of deaths ages 20-24 years
Number of deaths ages 5-9 years
Number of infant deaths
Number of infant deaths, female
Number of infant deaths, male

Number of neonatal deaths
Number of under-five deaths
Number of under-five deaths, female
Number of under-five deaths, male
Population ages 00-04, female
Population ages 00-04, female (% of female population)
Population ages 00-04, male
Population ages 00-04, male (% of male population)
Population ages 0-14 (% of total population)
Population ages 0-14, female
Population ages 0-14, female (% of female population)
Population ages 0-14, male
Population ages 0-14, male (% of male population)
Population ages 0-14, total
Population ages 05-09, female
Population ages 05-09, female (% of female population)
Population ages 05-09, male
Population ages 05-09, male (% of male population)
Population ages 10-14, female
Population ages 10-14, female (% of female population)
Population ages 10-14, male
Population ages 10-14, male (% of male population)
Population ages 15-19, female
Population ages 15-19, female (% of female population)
Population ages 15-19, male
Population ages 15-19, male (% of male population)
Population ages 15-64 (% of total population)
Population ages 15-64, female
Population ages 15-64, female (% of female population)
Population ages 15-64, male
Population ages 15-64, male (% of male population)
Population ages 15-64, total
Population ages 20-24, female
Population ages 20-24, female (% of female population)
Population ages 20-24, male
Population ages 20-24, male (% of male population)
Population ages 25-29, female
Population ages 25-29, female (% of female population)
Population ages 25-29, male
Population ages 25-29, male (% of male population)
Population ages 30-34, female
Population ages 30-34, female (% of female population)
Population ages 30-34, male
Population ages 30-34, male (% of male population)
Population ages 35-39, female
Population ages 35-39, female (% of female population)

Population ages 35-39, male
Population ages 35-39, male (% of male population)
Population ages 40-44, female
Population ages 40-44, female (% of female population)
Population ages 40-44, male
Population ages 40-44, male (% of male population)
Population ages 45-49, female
Population ages 45-49, female (% of female population)
Population ages 45-49, male
Population ages 45-49, male (% of male population)
Population ages 50-54, female
Population ages 50-54, female (% of female population)
Population ages 50-54, male
Population ages 50-54, male (% of male population)
Population ages 55-59, female
Population ages 55-59, female (% of female population)
Population ages 55-59, male
Population ages 55-59, male (% of male population)
Population ages 60-64, female
Population ages 60-64, female (% of female population)
Population ages 60-64, male
Population ages 60-64, male (% of male population)
Population ages 65 and above (% of total population)
Population ages 65 and above, female
Population ages 65 and above, female (% of female population)
Population ages 65 and above, male
Population ages 65 and above, male (% of male population)
Population ages 65 and above, total
Population ages 65-69, female
Population ages 65-69, female (% of female population)
Population ages 65-69, male
Population ages 65-69, male (% of male population)
Population ages 70-74, female
Population ages 70-74, female (% of female population)
Population ages 70-74, male
Population ages 70-74, male (% of male population)
Population ages 75-79, female
Population ages 75-79, female (% of female population)
Population ages 75-79, male
Population ages 75-79, male (% of male population)
Population ages 80 and above, female
Population ages 80 and above, male
Population ages 80 and above, male (% of male population)
Population ages 80 and older, female (% of female population)
Population growth (annual %)
Population, female

Population, female (% of total population)
Population, male
Population, male (% of total population)
Population, total
Probability of dying among adolescents ages 10-14 years (per 1,000)
Probability of dying among adolescents ages 15-19 years (per 1,000)
Probability of dying among children ages 5-9 years (per 1,000)
Probability of dying among youth ages 20-24 years (per 1,000)
Rural population
Rural population (% of total population)
Rural population growth (annual %)
Sex ratio at birth (male births per female births)
Urban population
Urban population (% of total population)
Urban population growth (annual %)

Appendix B – Environmental, Social and Governance Data

Archive file contains six csv files

ESGCountry.csv
ESGCountry-Series.csv
ESGData.csv
ESGFootnote.csv
ESGSeries.csv
ESGSeries-Time.csv

List of indicators in the file ESGData.csv

Access to clean fuels and technologies for cooking (% of population)
Access to electricity (% of population)
Adjusted savings: natural resources depletion (% of GNI)
Adjusted savings: net forest depletion (% of GNI)
Agricultural land (% of land area)
Agriculture, forestry, and fishing, value added (% of GDP)
Annual freshwater withdrawals, total (% of internal resources)
Annualized average growth rate in per capita real survey mean consumption or income, total population (%)
Cause of death, by communicable diseases and maternal, prenatal and nutrition conditions (% of total)
Children in employment, total (% of children ages 7-14)
CO2 emissions (metric tons per capita)
Control of Corruption: Estimate
Cooling Degree Days (projected change in number of degree Celsius)
Droughts, floods, extreme temperatures (% of population, average 1990-2009)
Ease of doing business rank (1=most business-friendly regulations)
Electricity production from coal sources (% of total)
Energy imports, net (% of energy use)
Energy intensity level of primary energy (MJ/\$2011 PPP GDP)
Energy use (kg of oil equivalent per capita)
Fertility rate, total (births per woman)
Food production index (2014-2016 = 100)
Forest area (% of land area)
Fossil fuel energy consumption (% of total)
GDP growth (annual %)
GHG net emissions/removals by LUCF (Mt of CO2 equivalent)
Gini index (World Bank estimate)
Government Effectiveness: Estimate
Government expenditure on education, total (% of government expenditure)
Heat Index 35 (projected change in days)
Hospital beds (per 1,000 people)
Income share held by lowest 20%

Individuals using the Internet (% of population)
Labor force participation rate, total (% of total population ages 15-64) (modeled ILO estimate)
Life expectancy at birth, total (years)
Literacy rate, adult total (% of people ages 15 and above)
Mammal species, threatened
Maximum 5-day Rainfall, 25-year Return Level (projected change in mm)
Mean Drought Index (projected change, unitless)
Methane emissions (metric tons of CO2 equivalent per capita)
Mortality rate, under-5 (per 1,000 live births)
Net migration
Nitrous oxide emissions (metric tons of CO2 equivalent per capita)
Patent applications, residents
People using safely managed drinking water services (% of population)
People using safely managed sanitation services (% of population)
PM2.5 air pollution, mean annual exposure (micrograms per cubic meter)
Political Stability and Absence of Violence/Terrorism: Estimate
Population ages 65 and above (% of total population)
Population density (people per sq. km of land area)
Poverty headcount ratio at national poverty lines (% of population)
Prevalence of overweight (% of adults)
Prevalence of undernourishment (% of population)
Proportion of seats held by women in national parliaments (%)
Ratio of female to male labor force participation rate (%) (modeled ILO estimate)
Regulatory Quality: Estimate
Renewable electricity output (% of total electricity output)
Renewable energy consumption (% of total final energy consumption)
Research and development expenditure (% of GDP)
Rule of Law: Estimate
School enrollment, primary (% gross)
School enrollment, primary and secondary (gross), gender parity index (GPI)
Scientific and technical journal articles
Strength of legal rights index (0=weak to 12=strong)
Terrestrial and marine protected areas (% of total territorial area)
Unemployment, total (% of total labor force) (modeled ILO estimate)
Unmet need for contraception (% of married women ages 15-49)
Voice and Accountability: Estimate