

Data Table Schema

tobacco_production

Tobacco production in metric tons and monetary value. Originally sourced from the United Nations Industrial Commodity Statistics Database, which provides annual statistics on the production of major industrial commodities by country.

1,207 rows & 5 columns. Size: <1MB. Source: [UNdata](#)

column	type	Description
Country or Area	STRING	Country or area
Year	INTEGER	Year of the measurement
Unit	STRING	Unit of the measurement
Value	NUMERIC	Value of the measurement
Footnotes	INTEGER	ID of the footnote. Look up using the tobacco_production_footnotes.txt file

tobacco_use_ww

The percentage of the population aged 15 years and over who currently use any tobacco product (smoked and/or smokeless tobacco) on a daily or non-daily basis. Tobacco products include cigarettes, pipes, cigars, cigarillos, water pipes (hookah, shisha), bidis, kretek, heated tobacco products, and all forms of smokeless (oral and nasal) tobacco. Tobacco products exclude e-cigarettes (which do not contain tobacco), “e-cigars”, “e-hookahs”, JUUL and “e-pipes”.

4,023 rows & 7 columns. Size: <1MB. Source: [WHO](#)

column	type	Description
ParentLocationCode	STRING	Parent location's code
ParentLocation	STRING	Parent location
SpatialDimValueCode	STRING	Location's code
Location	STRING	Location
Year	INTEGER	Year of measurement
Gender	STRING	Gender
Value	NUMBER	Value

tobacco_use_us

State Tobacco Activities Tracking and Evaluation (STATE) System. BRFSS Survey Data. The BRFSS is a continuous, state-based surveillance system that collects information about modifiable risk factors for chronic diseases and other leading causes of death. The data for the STATE System were extracted from the annual BRFSS surveys from participating states. Tobacco topics included are cigarette and e-cigarette use prevalence by demographics, cigarette and e-cigarette use frequency, and quit attempts.

111,048 rows & 21 columns. Size: ~30MB. Source: [CDC](#)

column	type	Description
Year	STRING	Year (or range) of the observation
LocationAbbr	STRING	Abbreviation of the location
LocationDesc	STRING	Description of the location
Topic	STRING	Description of the topic
Question	STRING	Question asked
Response	STRING	Type of response (every day, some days, ...)
Break_Out	STRING	Grouping (Male, 65+, 18-24, Hispanic, ...)
Break_Out_Category	STRING	Grouping category (age group, education, gender, ...)
Sample_Size	NUMERIC	Sample size
Data_Value	NUMERIC	The measurement
Confidence_limit_Low	NUMERIC	Low confidence limit
Confidence_limit_High	NUMERIC	High confidence limit
Data_Value_Unit	STRING	Unit of the data value (ex: %)
Data_Value_Type	STRING	Type of the data value (ex: percentage)
Data_Value_Footnote	STRING	Possible observation regarding this observation
LocationID	NUMERIC	Unique Identifier for the location
BreakoutID	STRING	Unique Identifier for the breakout
BreakOutCategoryID	STRING	Unique Identifier for the breakout category
QuestionID	STRING	Unique Identifier for the question
ResponseID	STRING	Unique Identifier for the response
GeoLocation	STRING	Coordinates

sales_per_day

Figures include manufactured cigarettes, as well as estimated number of hand-rolled cigarettes, per adult (ages 15+) per day. The number of cigarettes smoked per person per day for both males and females has been averaged across all years in which multiple estimates were provided in the ISS dataset for the United States, to arrive at one estimate for each year.

2,767 rows & 4 columns. Size: <1MB. Source: [ISS](#)

column	type	Description
Entity	STRING	Country
Code	STRING	Country code
Year	INTEGER	Year of measurement
amt	NUMERIC	Value of the measurement

death_rates_smoking_age

Death rates are measured as the number of early deaths due to smoking per 100,000 individuals in a given demographic group. Contains data from 1990 to 2017

6,468 rows & 9 columns. Size: <1MB. Source: [GHDX](#)

column	type	Description
Entity	STRING	Country
Code	STRING	Country code
Year	INTEGER	Year of measurement
All_ages	NUMERIC	Value of the measurement for all ages
Under_5	NUMERIC	Value of the measurement for ages under 5 years old
5_14	NUMERIC	Value of the measurement for ages from 5 to 14 years old
50_69	NUMERIC	Value of the measurement for ages from 50 to 69 years old
15_49	NUMERIC	Value of the measurement for ages from 15 to 49 years old
70_plus	NUMERIC	Value of the measurement for ages above 70 years old

stop_smoking

This dataset presents indicators that contribute to an individual to stop smoking. These contributions can be very direct, like offering government help, or less direct like increasing the tax and price of cigars and banning Tobacco advertisements.

In 2008, the World Health Organization (WHO) introduced the EMPOWER package to assist in the country-level implementation of six effective measures to reduce demand for tobacco products. Each measure reflects at least one provision of the WHO Framework Convention on Tobacco Control. This dataset covers 3 of those measures.

The letter E in POWER denotes the "Enforce bans on tobacco advertising" measure. Tobacco advertising, promotion and sponsorship can make smoking more socially acceptable, impede efforts to educate people about the hazards of tobacco use, and strengthen the tobacco industry's influence over media, sporting and entertainment businesses. WHO Framework Convention on Tobacco Control, Article 13 (Tobacco advertising, promotion and sponsorship) falls for comprehensive bans on tobacco advertising, promotion and sponsorship in accordance with each country's constitutional principles.

The letter O in MPOWER denotes the "Offer help to quit tobacco use" measure. Tobacco dependence treatment is primarily the responsibility of each country's health-care system. Individual cessation interventions have a significant impact on individual health and are extremely cost-effective compared with many other health system activities. People who quit smoking, regardless of their age, smoking history or health status, experience immediate and profound health benefits and can reduce most of the associated risks within a few years of quitting.

The letter R in MPOWER denotes the "Raise taxes on tobacco" measure.

774 rows & 7 columns. Size: <1MB. Source: [GHO](#)

column	type	Description
Entity	STRING	Country
Code	STRING	Country code
Year	INTEGER	Year of measurement
AvgCigarettePriceDollars	NUMERIC	Average price of a pack of 20 cigarettes, measured in international dollars. The average price is calculated based on the prices of three brands of cigarettes known to be most sold in the country. The average price is weighted by the market share of each of the three brands.
AvgTaxesAsPctCigarettePrice	NUMERIC	The sum of all taxes applicable on the weighted average price of the three most sold brands of cigarettes.

EnforceBansTobaccoAd	INTEGER	<p>Enforcement of bans on tobacco advertising are differentiated as follows:</p> <p>1 = Data not reported</p> <p>2 = Complete absence of ban, or ban that does not cover national television (TV), radio and print media</p> <p>3 = Ban on national TV, radio and print media only</p> <p>4 = Ban on national TV, radio and print media as well as on some but not all other forms of direct* and/or indirect** advertising</p> <p>5 = Ban on all forms of direct* and indirect** advertising.</p> <p>(See below for more information on direct and indirect advertisement)</p>
HelpToQuit	INTEGER	<p>Information from countries on the availability and non-availability of particular tobacco cessation aids is assessed to determine the comparative level of assistance countries provide to help tobacco users quit.</p> <p>The groupings for this indicator are:</p> <p>1 = Data not reported</p> <p>2 = None</p> <p>3 = NRT* and/or some cessation services** (neither cost-covered)</p> <p>4 = NRT* and/or some cessation services** (at least one of which is cost-covered)</p> <p>5 = National quitline, and both NRT* and some cessation services** cost-covered</p> <p>(See below for more information)</p>

Regarding the ***EnforceBansTobaccoAd*** field:

Examples of Direct advertising bans:

- national television and radio;
- local magazines and newspapers;
- billboards and outdoor advertising;
- point of sale.

Examples of Indirect advertising bans:

- free distribution of tobacco products in the mail or through other means;
- promotional discounts;
- non-tobacco products identified with tobacco brand names (brand extension);
- brand names of non-tobacco products used for tobacco products;
- appearance of tobacco products in television and/or films;
- sponsored events.

Regarding the ***HelpToQuit*** field:

* Nicotine replacement therapy.

** Smoking cessation support available in any of the following places: health clinics or other primary care facilities, hospitals, office of a health professional, the community.

us_chronic_resp_disease

Estimates for age-standardized mortality rates by county from chronic respiratory diseases. The estimates were generated using de-identified death records from the National Center for Health Statistics (NCHS); population counts from the U.S. Census Bureau, NCHS, and the Human Mortality Database; the cause list from the Global Burden of Disease Study (GBD); The table contains estimates for the age-standardized mortality rate (deaths per 100,000 population) for each sex and both sexes combined for each year in the period 1980-2014.

~3.4 million rows & 8 columns. Size: ~350MB. Source: [IHME](#)

column	type	Description
location_name	STRING	Location of the estimate
fips	STRING	The Federal Information Processing Standards (FIPS) code, a unique identifier for states and counties in the United States
cause_name	STRING	Cause of disease or injury of the estimate
sex	STRING	Gender for the estimate
year_id	STRING	Time period of estimate
mx	STRING	Posterior mean estimate
lower	STRING	2.5% percentile estimate
upper	STRING	97.5% percentile estimate