

Database files in IFs

The printable version is no longer supported and may have rendering errors. Please update your browser bookmarks and please use the default browser print function instead.

IFs Historical Database Files

IFs uses Microsoft Access files to store data and data dictionary (meta-data). All data files are in the "C:/My Documents/Users/Public/IFs/Data" folder. Data and related files are listed below:

- IFsHistSeries.Mdb is the largest and most frequently used IFs data file containing more than three thousand data tables each containing 186 rows (one row of data per country) and several columns (one column per year). The figure below shows data from an IFsHistSeries table

Country	FIPS_CODE	1991	1993	1995	1998	2000	2002	2004	Earliest	MostRecent
Afghanistan	AFG						0.34	0.53	0.34	0.53
Albania	ALB				0.86	0.57	0.8	1.23	0.86	1.23
Algeria	DZA	0.15	0.2	0.4	0.31	0.27	0.22	0.32	0.15	0.32
Angola	AGO				0.38	0.3	0.19	0.39	0.38	0.39
Argentina	ARG		0.79	0.6	0.94	1.07	0.63	0.63	0.79	0.63
Armenia	ARM				0.49	0.55	0.42	0.68	0.49	0.68
Australia	AUS				0.46	0.57	0.5	0.85	0.46	0.85
Austria	AUT			1.15	1.04	0.82	0.84	1.32	1.15	1.32

- DataDict.Mdb is the data dictionary file with a table containing one row of meta-data (e.g., definition, unit, source, last date of update) for each of the data tables in IFsHistSeries.Mdb

Variable	Table	Group	SubGroup	Series	Cr	Cohor	Definition
GDP2005PCPPPOLD	SeriesGDP2005PCPPPOLD	Economic	GDP per Capita	Yes	No	No	GDP per capita (constant 2005 US\$ PPP)
GDP2005PCPPPWDI2009	SeriesGDP2005PCPPPWDI2009	Economic	GDP per Capita	Yes	No	No	GDP per Capita, PPP (Constant 2005 Intern
GDP2008PCPPPWDI	SeriesGDP2008PCPPPWDI	Economic	Aggregate	Yes	No	No	GDP per capita at PPP (Current Internat
GDP95	SeriesGDP95	Economic	Aggregate	Yes	No	No	Gross Domestic Product
GDP98PPP	SeriesGDP98PPP	Economic	Aggregate	Yes	No	No	GDP (PPP)
GDPAngusMadison	SeriesGDPAngusMadison	Economic	Aggregate	Yes	No	No	Gross Domestic Product, Angus Maddison

- IFs.Mdb is the Microsoft Access file that contains several IFs data tables. One of these table - "Country Translation" - is required for automated IFs data import/update. Country Translation table maintains (and updates) a concordance list between country names used by IFs and data sources
- IFsWVSCohort.Mdb is the file that contains data from waves of World Value Survey, a global survey of cultural values conducted by University of Michigan.
- IFsDataImport.Mdb, is an MS Access database that holds the data series imported using IFs software's automated **single series 'import'** interface.
- IFsDataImportBatch.Mdb, is the Access database that houses the data series imported using IFs software's automated **batch** import interface.

IFs Data Table Naming Convention

Names of all data tables in IFsHistSeries.MDB start with the prefix "Series". The "Series" prefix is followed by an issue area prefix, e.g., "Ag" for agriculture or "Ed" for education. This second tier of prefix might be followed by additional prefixes (e.g., "EdSec" for secondary education) or might be absent altogether (e.g., in some of the earlier imports).

Data series names might also contain a suffix, the usual purpose of which is to differentiate among the sources for the same/similar series.

No spaces or symbols (other than %) are allowed in series name.

IFs DataDict Columns

The Datadict.mdb file serves as a reference for all series in the IFsHistSeries.mdb file. Every series in IFsHistSeries has an entry in DataDict containing all of the metadata on that series. The Data Dictionary lists each variable, the groups to which it belongs (e.g., Agriculture, Economics) its subgroup (e.g., Trade, Consumption), and additional identifying information. This information includes whether or not the data is a series (Yes/No), CoVaTra, Cohort. It also includes a definition of the variable, and a column for an extended definition provided by the data source. The data dictionary has columns identifying the years for which a series has data, the source of the data, the original source (e.g. a series may have been pulled from the FAO website, but may have originated as World Bank research.) and the source name of the series, and an identifier for which team member last updated the series and when. It also includes instructions on how data should be aggregated or disaggregated for provincial models (e.g., by population or GDP distribution). Some additional information is supplied that is used by the model such as whether a datum of 0 should be treated as a null or as a zero, if a series is used in the preprocessor, if it is compared to other forecasts, the number of decimal places to read, and any formulas applied to the data.

Column	Description	Example
Variable Name	Name of variable you assign (Automatically generated from import)	WSSJMPSanitationRural%Improved
Table	Automatically generated	SeriesWSSJMPSanitationRural%Improved
Group	Group or category you assign to the variable (Automatically generated from import)	Infrastructure, Water, Health
Subgroup	Subroup you assign to the variable (Automatically generated from import)	Sanitation
Series	Yes	Yes
CoVaTra	No	No
Cohort	No	No
Definition	Specific definition of the variable that will be displayed in IFs	Proportion of Rural population served with Improved Sanitation (%)
Extended Source Definition	Additional information about the variable if needed	Likely to ensure hygienic separation of human excreta from human contact.
Units	Specification of the units the variable is measured by	Percent;People;Degrees C
Currency	Currency of the units if they're monetary	USD;Euro;Lira
Years	Years that data set covers (Automatically generated but make sure its accurate)	2000-2015
Source	Source or website the data source is from	World Bank
Original Source	URL of data source	https://washdata.org
Notes	Any additional notes and the initials of the puller and vetter of the data	Created individual index;EB;JD
Last IFs Update	Automatically generated	10/30/2017

Aggregation	Tells the model how to aggregate country values to groups within the model. Will automatically generate if classified during import.	POP;GDP;LAND;AVG;SUM
Disaggregation	Tells the model how to disaggregate country values to subregional classification	GDP
TreatNullsAs0s	If checked, will treat nulls in the data set as zeroes	
Proprietary	Usually data series are non-proprietary	
Name in Source	Exact name of series in the source file and other nformation on how the data was pulled so someone can reimport it in the same way	World File;Sanitation;Sum of Rural Latrines, Septic Tanks, Sewer Connections
Used in Preprocessor	Filled automatically when script is run before consolidation. If checked, data set is used in preprocessor. Puller/Vetter not responsible for this.	
UsedInPreprocessorFileName	File name that is used in preprocessor that is filled automatically before consolidation. Puller/vetter not responsible for this.	
CompareOtherForecasts	If series is a forecast from another source, this will be checked	
Code in Source	Only used in batch pulls	
Decimal Places	Decimal places for data in data series	4
Country Concordance	Country concordance used to import data into IFs. (Will automatically generate)	IFs Country
Formula	Option during import to manipulate the data	*100 (to convert to %)

Retrieved from "https://pardeewiki.du.edu//index.php?title=Database_files_in_IFs&oldid=8695"

This page was last edited on 9 November 2017, at 18:46.