Changes made:

* Decomposed bird table
* Removed locality column from bird table
* Added constraints ~~(except for worldBankData constraint: FOREIGN KEY (indicatorCode) REFERENCES worldBankIndicators (indicatorCode), because it failed [RESOLVED])~~
* Renamed worldBankSeriesData to worldBankIndicators and renamed ‘seriesCode’ column in worldBankSeriesData to ‘indicatorCode’
* Identified discrepancy behind failed constraint, added the 2 missing rows into worldBankIndicators to address
* Found out that failed constraint is due to missing rows in worldBankIndicators that were taken out due to discrepancies in Indicator Name and have added those 2 rows into worldBankIndicators and added the foreign key constraint

# BEFORE

# view table DDL (before altering tables)

SHOW CREATE TABLE birdData;

CREATE TABLE `birdData` (

`id` varchar(25) DEFAULT NULL,

`genus` varchar(25) DEFAULT NULL,

`scientificName` varchar(100) DEFAULT NULL,

`vernacularName` varchar(100) DEFAULT NULL,

`longitudeDecimal` float DEFAULT NULL,

`latitudeDecimal` float DEFAULT NULL,

`country` varchar(25) DEFAULT NULL,

`locality` varchar(100) DEFAULT NULL,

`accessURI` varchar(150) DEFAULT NULL

)

SHOW CREATE TABLE worldBankData;

CREATE TABLE `worldBankData` (

`countryName` varchar(25) DEFAULT NULL,

`indicatorCode` varchar(25) DEFAULT NULL,

`year` int DEFAULT NULL,

`value` float DEFAULT NULL

)

SHOW CREATE TABLE worldBankSeriesData;

CREATE TABLE `worldBankSeriesData` (

`seriesCode` varchar(25) DEFAULT NULL,

`topic` text,

`indicatorName` text,

`longDefinition` text,

`unitOfMeasure` text,

`periodicity` varchar(25) DEFAULT NULL,

`basePeriod` varchar(25) DEFAULT NULL,

`otherNotes` text,

`aggregationMethod` text,

`limitations` text,

`notes` text,

`generalComments` text,

`source` text,

`statConceptMethod` text,

`devRelevance` text

)

# SQL statements

# create birdSpecies table

CREATE TABLE birdSpecies AS

SELECT DISTINCT scientificName, vernacularName, genus

FROM birdData;

ALTER TABLE birdSpecies

ADD PRIMARY KEY (scientificName);

MODIFY vernacularName varchar(100) NOT NULL,

MODIFY genus varchar(25) NOT NULL;

# only run the following query (drop columns from birdData) after creating birdSpecies table

# moving ‘genus’ and ‘vernacularName’ to different table for 3NF decomposition

# we don’t need ‘locality’ column

# and add constraints

ALTER TABLE birdData

DROP COLUMN locality,

DROP COLUMN genus,

DROP COLUMN vernacularName,

ADD PRIMARY KEY (id),

ADD FOREIGN KEY (scientificName) REFERENCES birdSpecies(scientificName),

MODIFY scientificName varchar(100) NOT NULL,

MODIFY country varchar(25) NOT NULL,

MODIFY accessURI varchar(150) NOT NULL;

# rename worldBankIndicators for clarity

# and add constraints

ALTER TABLE worldBankSeriesData

RENAME TO worldBankIndicators,

RENAME COLUMN seriesCode TO indicatorCode,

ADD PRIMARY KEY (indicatorCode),

MODIFY topic text NOT NULL,

MODIFY indicatorName text NOT NULL;

# add constraints - FAILED

~~ALTER TABLE worldBankData~~

~~ADD PRIMARY KEY (countryName, indicatorCode, year),~~

~~ADD FOREIGN KEY (indicatorCode) REFERENCES worldBankIndicators(indicatorCode),~~

~~MODIFY value float NOT NULL;~~

# try again without the failing constraint

ALTER TABLE worldBankData

ADD PRIMARY KEY (countryName, indicatorCode, year),

MODIFY value float NOT NULL;

# identified failed foreign key constraint (see Jupyter notebook with relevant name)

# to address the failed foreign key constraint

# after identifying the discrepancies between worldBankData and worldBankIndicators,

# fix the discrepancies:

# add missing row to worldBankIndicators

INSERT INTO worldBankIndicators(indicatorCode, topic, indicatorName, longDefinition, source)

VALUES ('DT.TDS.DPPF.XP.ZS', 'Economic Policy & Debt: External debt: Debt service', 'Present value of external debt (% of exports of goods, services and primary income)', 'Debt service, the sum of principal repayments and interest actually paid in currency, goods, or services, is expressed as a percentage of exports of goods and services--all transactions between residents of a country and the rest of the world involving a change of ownership from residents to nonresidents of general merchandise, net exports of goods under merchanting, nonmonetary gold, and services. This series differs from the standard debt to exports series in that it covers only long-term public and publicly guaranteed debt and repayments (repurchases and charges) to the IMF.', 'World Bank, International Debt Statistics.');

# check that this is correct before running it

# add other missing row to worldBankIndicators

INSERT INTO worldBankIndicators (indicatorCode, topic, indicatorName, longDefinition, periodicity, source, statConceptMethod, devRelevance)

VALUES ('DT.DOD.PVLX.EX.ZS', 'Economic Policy & Debt: External debt: Debt ratios & other items', 'Present value of external debt (% of exports of goods, services and primary income', 'Present value of debt is the sum of short-term external debt plus the discounted sum of total debt service payments due on public, publicly guaranteed, and private nonguaranteed long-term external debt over the life of existing loans. This calculation assumes that the PV of loans with a negative grant element is equal to the nominal value of the loan. The exports denominator is a three-year average.', 'Annual', 'World Bank, International Debt Statistics.', 'Data on external debt are gathered through the World Bank\'s Debtor Reporting System (DRS). Long term debt data are compiled using the countries report on public and publicly guaranteed borrowing on a loan-by-loan basis and private non guaranteed borrowing on an aggregate basis. These data are supplemented by information from major multilateral banks and official lending agencies in major creditor countries. Short-term debt data are gathered from the Quarterly External Debt Statistics (QEDS) database, jointly developed by the World Bank and the IMF and from creditors through the reporting systems of the Bank for International Settlements. Debt data are reported in the currency of repayment and compiled and published in U.S. dollars. End-of-period exchange rates are used for the compilation of stock figures (amount of debt outstanding), and projected debt service and annual average exchange rates are used for the flows. Exchange rates are taken from the IMF\'s International Financial Statistics. Debt repayable in multiple currencies, goods, or services and debt with a provision for maintenance of the value of the currency of repayment are shown at book value.', 'External debt is that part of the total debt in a country that is owed to creditors outside the country. The debtors can be the government, corporations or private households. The debt includes money owed to private commercial banks, other governments, or international financial institutions.

\n

\nExternal indebtedness affects a country\'s creditworthiness and investor perceptions. Nonreporting countries might have outstanding debt with the World Bank, other international financial institutions, or private creditors. Total debt service is contrasted with countries\' ability to obtain foreign exchange through exports of goods, services, primary income, and workers\' remittances.

\n

\nDebt ratios are used to assess the sustainability of a country\'s debt service obligations, but no absolute rules determine what values are too high. Empirical analysis of developing countries\' experience and debt service performance shows that debt service difficulties become increasingly likely when the present value of debt reaches 200 percent of exports. Still, what constitutes a sustainable debt burden varies by country. Countries with fast-growing economies and exports are likely to be able to sustain higher debt levels. Various indicators determine a sustainable level of external debt, including:

\n

\na) debt to GDP ratio

\nb) foreign debt to exports ratio

\nc) government debt to current fiscal revenue ratio

\nd) share of foreign debt

\ne) short-term debt

\nf) concessional debt in the total debt stock');

# now try to add the foreign key constraint again:

ALTER TABLE worldBankData

ADD FOREIGN KEY (indicatorCode) REFERENCES worldBankIndicators(indicatorCode);

# AFTER

# view table DDL after altering tables

SHOW CREATE TABLE birdData;

CREATE TABLE `birdData` (

`id` varchar(25) NOT NULL,

`scientificName` varchar(100) NOT NULL,

`longitudeDecimal` float DEFAULT NULL,

`latitudeDecimal` float DEFAULT NULL,

`country` varchar(25) NOT NULL,

`accessURI` varchar(150) NOT NULL,

PRIMARY KEY (`id`),

KEY `scientificName` (`scientificName`),

CONSTRAINT `birdData\_ibfk\_1` FOREIGN KEY (`scientificName`) REFERENCES `birdSpecies` (`scientificName`)

)

SHOW CREATE TABLE birdSpecies;

CREATE TABLE `birdSpecies` (

`scientificName` varchar(100) NOT NULL,

`vernacularName` varchar(100) NOT NULL,

`genus` varchar(25) NOT NULL,

PRIMARY KEY (`scientificName`)

)

SHOW CREATE TABLE worldBankData;

CREATE TABLE `worldBankData` (

`countryName` varchar(25) NOT NULL,

`indicatorCode` varchar(25) NOT NULL,

`year` int NOT NULL,

`value` float NOT NULL,

PRIMARY KEY (`countryName`,`indicatorCode`,`year`),

KEY `indicatorCode` (`indicatorCode`),

CONSTRAINT `worldBankData\_ibfk\_1` FOREIGN KEY (`indicatorCode`) REFERENCES `worldBankIndicators` (`indicatorCode`)

)

SHOW CREATE TABLE worldBankIndicators;

CREATE TABLE `worldBankIndicators` (

`indicatorCode` varchar(25) NOT NULL,

`topic` text NOT NULL,

`indicatorName` text NOT NULL,

`longDefinition` text,

`unitOfMeasure` text,

`periodicity` varchar(25) DEFAULT NULL,

`basePeriod` varchar(25) DEFAULT NULL,

`otherNotes` text,

`aggregationMethod` text,

`limitations` text,

`notes` text,

`generalComments` text,

`source` text,

`statConceptMethod` text,

`devRelevance` text,

PRIMARY KEY (`indicatorCode`)

)