Risk Ranking Tool User's Guide

April 17, 2009
RTI Number 0211460.001

Submitted to:
U.S. Food and Drug Administration
Center for Food Safety and Applied Nutrition
5100 Paint Branch Parkway
College Park, Maryland 20740-3835

Submitted by:
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709-2194

1.0 Introduction

The U.S. Food and Drug Administration (FDA's) Risk Ranking Tool (RRT) application was developed based on a simple, transparent, risk ranking algorithm to rank order the priority pathogen-commodity combinations. The methods for this procedure are based on the Relative Ranking procedure of determining the severity of the health effect for each hazard as well as the likelihood that the hazard occurs in fresh produce and the likelihood that it will cause illness or death. The methodology used in the RRT expands on this approach by also including the probability of consumption, contamination, infectious dose, growth potential, hospitalizations, deaths, and if there is an epidemiological link and susceptible populations. The shelf life of the commodity in the risk ranking has also been added.

2.0 Underlying data

The underlying data the rankings are based on include information is eleven different data categories. The agent data categories include infectious dose and susceptible population. Hospitalization, death, epidemiological link, and epidemiological multipliers are the health data categories. The final data type, production/processing, includes data categories contamination, consumption, and shelf life and growth potential.

Data sources utilized to compile the dataset include:

- All of the reported foodborne outbreaks of confirmed etiology associated with fresh produce from 1996 to 2006 were compiled from the CDC's reports on foodborne outbreaks (http://www.cdc.gov/foodborneoutbreaks/outbreak_data.htm) along with the last ten years of the MMWR and peer reviewed journals articles that were found during the literature search on PubMed.

- Peer reviewed literature found through the PubMed search was used to compile prevalence data for pathogens found on fresh produce.

- The National Health and Nutrition Examination Survey (NHANES) database was used to generate consumption data for the fresh produce items found to be associated with outbreaks in our epidemiological database. The NHANES database is a three day dietary recall database compiled by the Centers for Disease Control (http://www.cdc.gov/nchs/nhanes.htm).

A data dictionary and entity relationship documents for the risk ranking tool can be found in Appendix A.

3.0 Risk Ranking Tool

The RRT facilitates rapid ranking of the priority list based on user defined criteria which includes refining the default bins for each data category with numerical criteria and entering weights for the data category. The tool provides a basic mechanism to semi-quantitatively compare the relative importance of one

pathogen-commodity combination to others, taking into account the current body of evidence related to the production and processing of those products, and the disease-causing characteristics of the agents. The risk-ranking summary report generated provides the list of pathogen-commodity combination ordered by total risk in descending order as well as inputs used to generate the list.

The RRT is a Microsoft (MS) Access database application and MS Access 2000, 2003, or 2007 is required to run the tool. Once the database file is downloaded, you can open it by double clicking on the file. When the database opens, main screen, Figure 1, of the RRT will be visible.

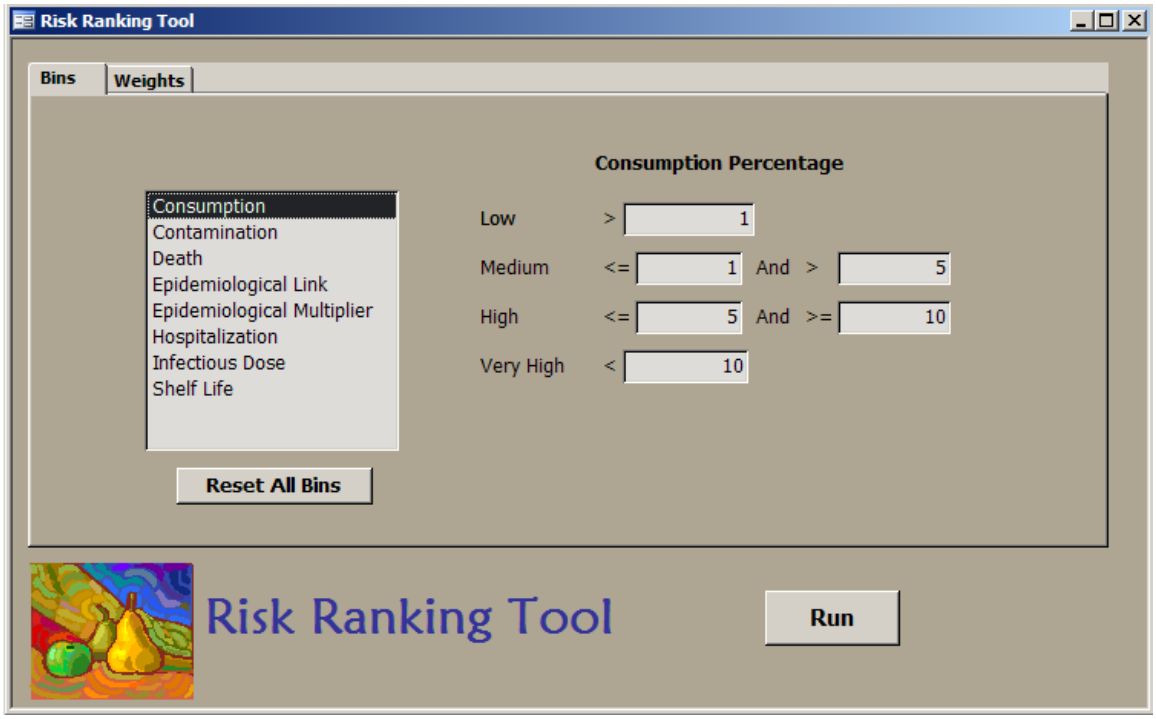

Figure 1. Risk Ranking Tool Main Screen

On the left hand side of the screen is a list of the data categories with numeric criteria. On the right hand side are the bins for the select data category, in this case Consumption. You can look at the bins for each data category by clicking on the data category name in the list. The bin ranges for each data category can be changed by overwriting the numbers in the boxes. Note that the application requires the ranges to be continuous, so for example, if you enter in 2 in the "Low" bin, when you exit the textbox the application will automatically put 2 in the lower end of the range of the "Medium" bin. To reset all the data category bins to the default values, you can click the "Reset All Bins" button located under the data category list.

When you have completed any bin range adjustments, select the "Weights" tab to assign weights to each of the data categories. Figure 2 shows the weights screen.

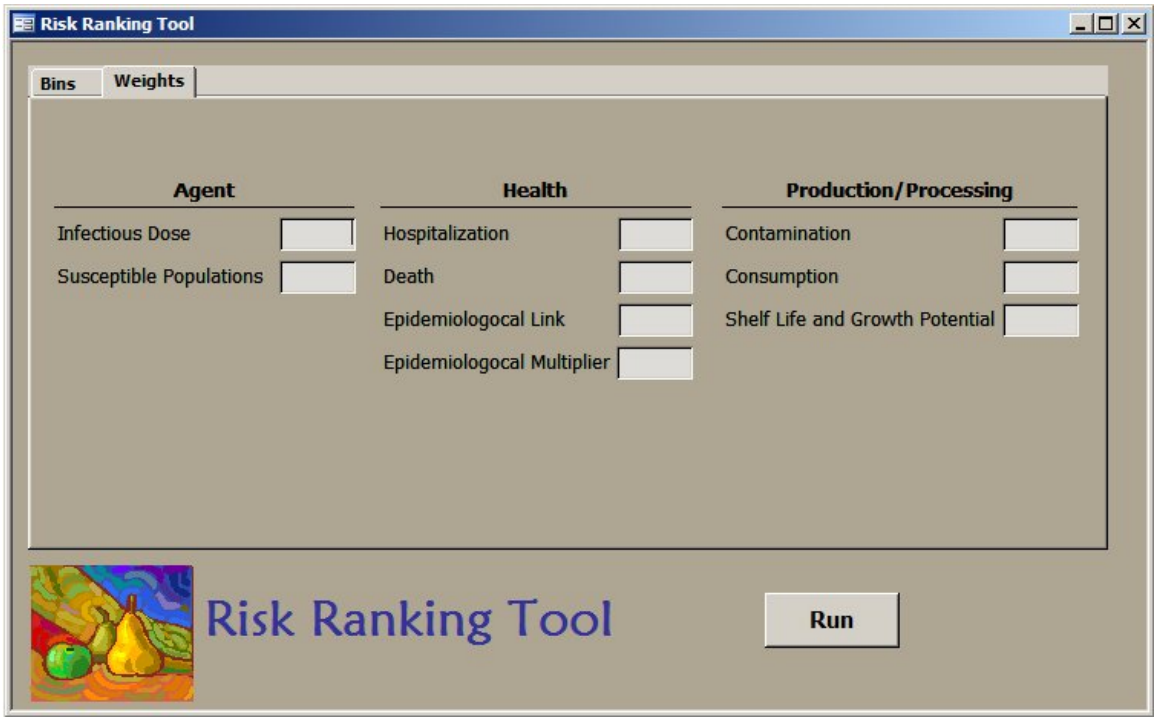

Figure 2. Risk Ranking Tool Weights Screen

The data categories are arranged by the data types, agent, health, and production/processing. Enter in weights for each data category. To run the ranking, all the data categories need to have a weight, to exclude a data category from the run, enter 0 in the textbox.

Once the bins and weights are set select the "Run" button to calculate the risk ranking. Once the application completes the ranking, a report with the results will be displayed. The results are pathogen – commodity pairs ranked in descending order of the score. Figure 3 shows the risk ranking results report.

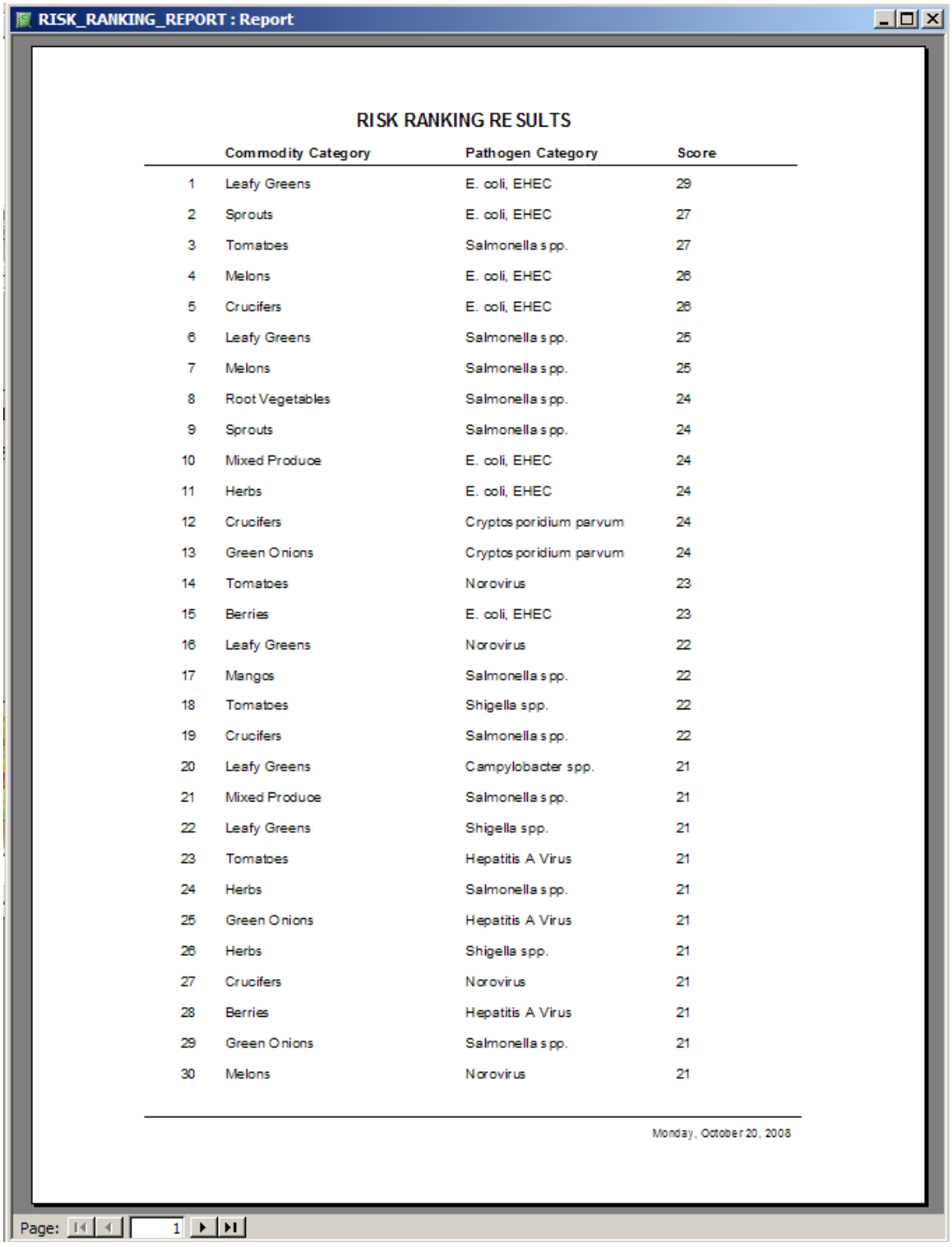

Figure 3. Risk Ranking Results Report

Figure 4 shows the user inputs for the risk ranking calculation at the end of the report.

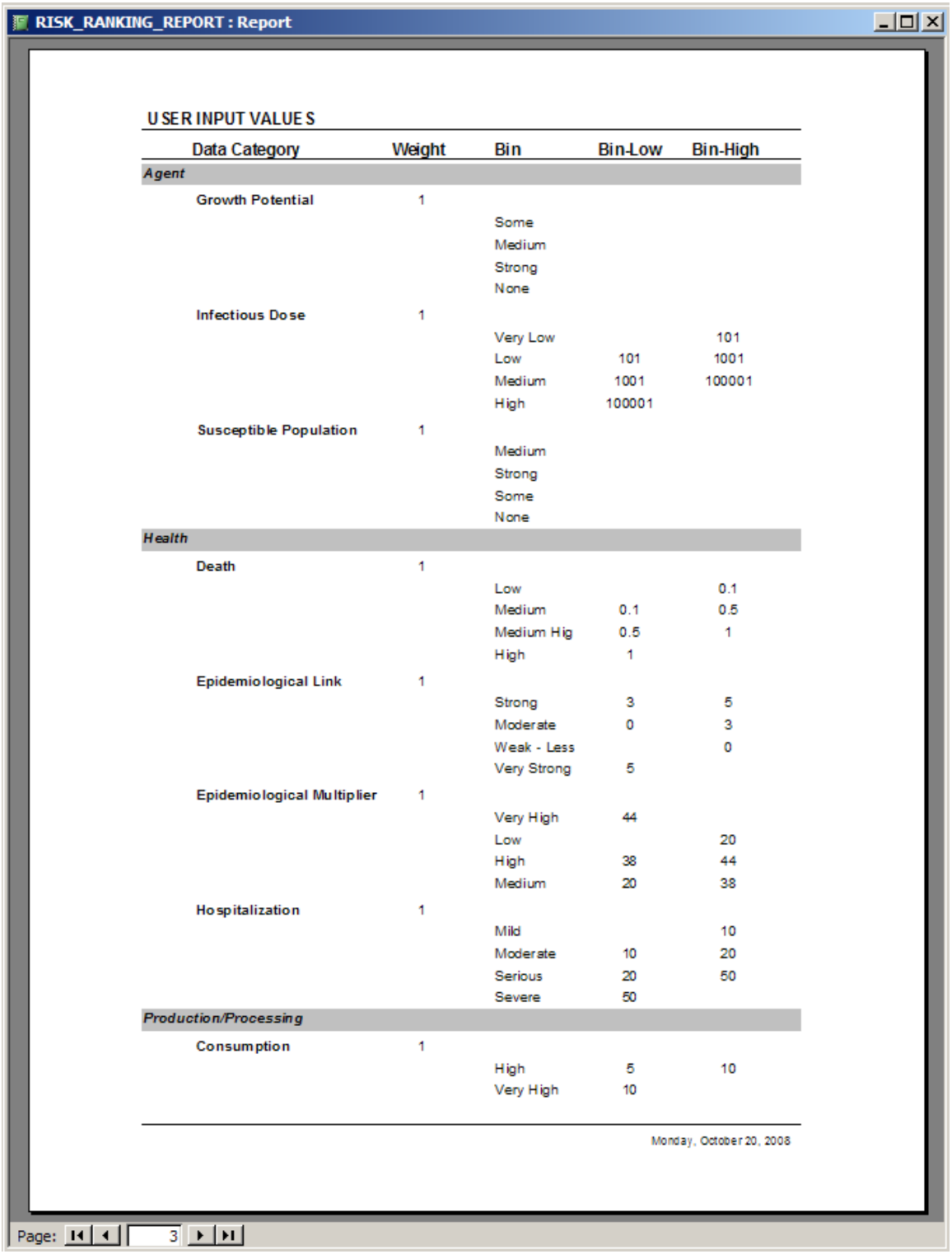

Figure 4. User Inputs at the end of the Risk Ranking Results Report

Appendix A

Risk Ranking Tool Data Dictionary

Risk Ranking Tool Data Dictionary
August 2008

BINS

Contains the bin descriptions and high and low values for each data category

Field Name	Description	Data Type	Not Null	Constraint	Comment
DATA_CATEGORY_ID	Unique identifier assigned to a data category	INTEGER	Y	PK, FK	FK – DATA_CATEGORY_LUT
BIN_NUMBER	Identifies the bin number (1-4)	INTEGER	Y	PK	
BIN_RANK	Rank/Risk assigned to the bin number (1-4)	INTEGER	Y		
BIN	Description of the bin for each data category	TEXT(255)	Y		
BIN_RANGE	Description of the bin range	TEXT(255)	Y		
BIN_LOW_DEFAULT	Default value for the low end of the bin	NUMBER	N		
BIN_HIGH_DEFAULT	Default value for the high end of the bin	NUMBER	N		
BIN_LOW_USER	User defined value for the low end of the bin, defaulted initially to the BIN_LOW_DEFAULT value	NUMBER	N		
BIN_HIGH_USER	User defined value for the high end of the bin, defaulted initially to the BIN_HIGH_DEFAULT value	NUMBER	N		

COMMODITY_CATEGORY_LUT

Look up table for commodity categories. This is the level the ranking is done on.

Field Name	Description	Data Type	Not Null	Constraint	Comment
COMMODITY_CATEGORY_ID	Unique identifier for the commodity category	INTEGER	Y	PK	
COMMODITY_CATEGORY	Name of the commodity category to which commodities are assigned. The tool ranks at this level.	TEXT(50)	Y		

COMMODITY_ LUT

Look up table for the individual commodities.

Field Name	Description	Data Type	Not Null	Constraint	Comment
COMMODITY_ID	Unique identifier for the commodities	INTEGER	Y	PK	
COMMODITY	Name of the individual commodity	TEXT(100)	Y		
COMMODITY_CATEGORY_ID	Identifier of the commodity category to which the commodity belongs	INTEGER	Y	FK	FK – COMMODITY_CATEGORY_ LUT

CONSUMPTION

Contains consumption data summarized to the commodity category level. (Commodity Category)

Field Name	Description	Data Type	Not Null	Constraint	Comment
COMMODITY_CATEGORY_ID	Identifier for the commodity category	INTEGER	Y	PK	FK – COMMODITY_CATEGORY_ LUT
AVG_POP_CONSUMING_PERCENT	Average of the population consuming percentages at the commodity category level. Calculated from the raw consumption data collected at commodity and/or commodity category level	NUMBER	Y		
USER_BIN_RANK	Rank assigned to this commodity category for consumption based on where the value in AVG_POP_CONSUMING_PERCENT falls in the user defined bins	INTEGER	Y		

CONSUMPTION_RAW

Contains consumption data summarized to the commodity category level. (Commodity Category)

Field Name	Description	Data Type	Not Null	Constraint	Comment
COMMODITY_ID	Identifier for the commodity	INTEGER	N		FK – COMMODITY_ LUT
COMMODITY_CATEGORY_ID	Identifier for the commodity category	INTEGER	Y		FK – COMMODITY_CATEGORY_ LUT
MEAN_GRAMS	Average grams of commodity consumed	NUMBER	Y		
POP_CONSUMING_PERCENT	Percentage of the population consuming the commodity/commodity category	NUMBER	Y		

CONTAMINATION

Contains consumption data summarized to the commodity category level. (Commodity Category)

Field Name	Description	Data Type	Not Null	Constraint	Comment
PATHOGEN_CATEGORY_ID	Identifier for the pathogen category	INTEGER	Y	PK	FK – PATHOGEN_CATEGORY_L UT
COMMODITY_CATEGORY_ID	Identifier for the commodity category	INTEGER	Y	PK	FK – COMMODITY_CATEGORY_ LUT
PREVALENCE_RANGE_PERCE NT	Percentage range of pathogen category - commodity category prevalence	TEXT(255)	Y		
AVG_PREVALENCE_PERCENT	Average prevalence percent based on prevalence range	NUMBER	Y		
NUMBER_STUDIES	Number of studies referencing this pathogen category - commodity category	INTEGER	Y		
USER_BIN_RANK	Rank assigned to this commodity category for consumption based on where the value in AVG_POP_CONSUMING_PERCENT falls in the user defined bins	INTEGER	Y		

DATA_CATEGORY_LUT
Look up table for the data categories.

Field Name	Description	Data Type	Not Null	Constraint	Comment
DATA_CATEGORY_ID	Identifier for the data category	INTEGER	Y	PK	
DATA_CATEGORY	Name of data category	TEXT(50)	Y		
DATA_TYPE_ID	Identifier for the data type	INTEGER	Y		FK – DATA_TYPE_LUT
NUMERIC_CRITERIA_FLAG	Flag to track whether the evaluation criteria for the bins is numeric or narrative	BOOLEAN	Y		
BIN_DESC	Description of the bin (including unit where applicable) for each data category	TEXT(100)	Y		
USER_WEIGHT	Weight assigned by the user for the data category	NUMBER	Y		

DATA_TYPE_LUT
Look up table for the data types.

Field Name	Description	Data Type	Not Null	Constraint	Comment
DATA_TYPE_ID	Unique identifier for the data type	INTEGER	Y	PK	
DATA_TYPE	Name of the data type	TEXT(50)	Y		

DEATH
Contains death data. (Pathogen Category)

Field Name	Description	Data Type	Not Null	Constraint	Comment
PATHOGEN_CATEGORY_ID	Identifier for the pathogen category	INTEGER	Y	PK	FK – COMMODITY_CATEGORY_LUT
DEATH_PERCENT	Death percentage for the pathogen category	NUMBER	Y		
USER_BIN_RANK	Rank assigned to this pathogen category for death based on where the value in DEATH_PERCENT falls in the user defined bins	INTEGER	Y		

EPIDEMIOLOGICAL_LINK

Contains epidemiological link data. (Pathogen Category and Commodity Category)

Field Name	Description	Data Type	Not Null	Constraint	Comment
PATHOGEN_CATEGORY_ID	Identifier for the pathogen category	INTEGER	Y	PK	FK – PATHOGEN_CATEGORY_LUT
COMMODITY_CATEGORY_ID	Identifier for the commodity category	INTEGER	Y	PK	FK – COMMODITY_CATEGORY_LUT
OUTBREAKS	Number of outbreaks for this pathogen category - commodity category pair	NUMBER	Y		
TOTAL_CASES	Number of case for this pathogen category - commodity category pair	NUMBER	Y		
USER_BIN_RANK	Rank assigned to this pathogen category - commodity category pair for epidemiological link based on where the value in AVG_NUMBER_PER_OUTBREAK falls in the user defined bins	INTEGER	Y		

EPIDEMIOLOGICAL_MULTIPLIERS

Contains Epidemiological Multiplier data. (Pathogen Category and Commodity Category)

Field Name	Description	Data Type	Not Null	Constraint	Comment
PATHOGEN_CATEGORY_ID	Identifier for the pathogen category	INTEGER	Y	PK	FK – PATHOGEN_CATEGORY_LUT
COMMODITY_CATEGORY_ID	Identifier for the commodity category	INTEGER	Y	PK	FK – COMMODITY_CATEGORY_LUT
MULTIPLIER	Multiplier assigned to each pathogen category - commodity category pair	NUMBER	Y		
USER_BIN_RANK	Rank assigned to this pathogen category - commodity category for epidemiological multiplier based on where the value in MULTIPLIER falls in the user defined bins	INTEGER	Y		

GROWTH_POTENTIAL

Contains growth potential data. (Pathogen Category and Commodity Category)

Field Name	Description	Data Type	Not Null	Constraint	Comment
PATHOGEN_CATEGORY_ID	Identifier for the pathogen category	INTEGER	Y	PK	FK – PATHOGEN_CATEGORY_LUT
COMMODITY_CATEGORY_ID	Identifier for the commodity category	INTEGER	Y	PK	FK – COMMODITY_CATEGORY_LUT
GROWTH_POTENTIAL_COMMENTS	Growth potential comment for this commodity category-pathogen category pair	TEXT(255)	Y		
DEFAULT_BIN	Default bin value	TEXT(50)	Y		
USER_BIN	User defined bin value, default value is DEFAULT_BIN, can be manually changed by the user in this table	TEXT(50)	Y		

USER_BIN_RANK	Rank assigned to the commodity category-pathogen category combination for growth potential based on where the value in USER_BIN falls in the user defined bins	INTEGER	Y		

HOSPITALIZATION

Contains hospitalization data. (Pathogen Category)

Field Name	Description	Data Type	Not Null	Constraint	Comment
PATHOGEN_CATEGORY_ID	Identifier for the pathogen category	INTEGER	Y	PK	FK – COMMODITY_CATEGORY_LUT
HOSPITALIZATION_PERCENT	Hospitalization percentage for the pathogen category	NUMBER	Y		
USER_BIN_RANK	Rank assigned to this pathogen category for hospitalization based on where the value in HOSPITALIZATION_ PERCENT falls in the user defined bins	INTEGER	Y		

INFECTIOUS_DOSE

Contains infectious dose data. (Pathogen Category)

Field Name	Description	Data Type	Not Null	Constraint	Comment
PATHOGEN_CATEGORY_ID	Identifier for the pathogen category	INTEGER	Y	PK	FK – COMMODITY_CATEGORY_LUT
INFECTIOUS_DOSE_RANGE	Range of the number of organism in an infectious dose of this pathogen category	TEXT(255)	Y		
INFECTIOUS_DOSE_NUM_ORGANISMS	Number of organism in an infectious dose of this pathogen category	NUMBER	Y		
USER_BIN_RANK	Rank assigned to this pathogen category for infectious dose based on where the value in INFECTIOUS_DOSE_NUM_ ORGANISMS falls in the user defined bins	INTEGER	Y		

PATHOGEN_CATEGORY_LUT

Look up table for pathogen categories. This is the level the ranking is done on. Contains the mapping to the pathogen type level.

Field Name	Description	Data Type	Not Null	Constraint	Comment
PATHOGEN_CATEGORY_ID	Unique identifier for the pathogen category	INTEGER	Y	PK	
PATHOGEN _CATEGORY	Name of the pathogen category to which pathogens are assigned. The tool ranks at this level.	TEXT(50)	Y		

PATHOGEN_ LUT

Look up table for the individual pathogens. Contains the mapping to the pathogen category level.

Field Name	Description	Data Type	Not Null	Constraint	Comment
PATHOGEN _ID	Unique identifier for the pathogens	INTEGER	Y	PK	
PATHOGEN	Name of the individual pathogen	TEXT(100)	Y		
PATHOGEN _CATEGORY_ID	Identifier of the pathogen category to which the pathogen belongs	INTEGER	Y	FK	FK – PATHOGEN _CATEGORY_LUT

PATHOGEN_TYPE_LUT

Look up table for the pathogen types.

Field Name	Description	Data Type	Not Null	Constraint	Comment
PATHOGEN _TYPE_ID	Unique identifier for the pathogen type	INTEGER	Y	PK	
PATHOGEN _TYPE	Name of the pathogen type	TEXT(50)	Y		

RANK_RESULTS

Contains results of the risk ranking run. This table is the default for the pathogen category - commodity category pairs that will be ranked.

Field Name	Description	Data Type	Not Null	Constraint	Comment
PATHOGEN_CATEGORY_ID	Identifier for the pathogen category	INTEGER	Y	PK	FK – PATHOGEN_CATEGORY_LUT
COMMODITY_CATEGORY_ID	Identifier for the commodity category	INTEGER	Y	PK	FK – COMMODITY_CATEGORY_LUT
RANK_TOTAL	Ranking total calculated based on the bins and weights provided by the user	NUMBER	Y		

SHELF_LIFE

Contains shelf life data summarized to the commodity category level. (Commodity Category)

Field Name	Description	Data Type	Not Null	Constraint	Comment
COMMODITY_CATEGORY_ID	Identifier for the commodity category	INTEGER	Y	PK	FK – COMMODITY_CATEGORY_LUT
AVG_SHELF_LIFE_DAYS	Shelf life of the commodity category. Calculated from the raw shelf life data collected at commodity and/or commodity category level	NUMBER	Y		
USER_BIN_RANK	Rank assigned to this commodity category for shelf life based on where the value in AVG_SHELF_LIFE_DAYS falls in the user defined bins	INTEGER	Y		

SHELF_LIFE_GROWTH_POTENTIAL_COMBINED

Contains growth potential data and shelf life combined user bin ranking. (Pathogen Category and Commodity Category)

Field Name	Description	Data Type	Not Null	Constraint	Comment
PATHOGEN_CATEGORY_ID	Identifier for the pathogen category	INTEGER	Y	PK	FK – PATHOGEN_CATEGORY_LUT
COMMODITY_CATEGORY_ID	Identifier for the commodity category	INTEGER	Y	PK	FK – COMMODITY_CATEGORY_LUT
SUM_SL_GP_RANK	Sum of the shelf life and growth potential rankings	NUMBER	Y		
USER_BIN_RANK	Rank assigned to this commodity category for shelf life based on where the value in AVG_SHELF_LIFE_DAYS falls in the user defined bins	INTEGER	Y		

SHELF_LIFE_RAW

Contains shelf life data as collected.

Field Name	Description	Data Type	Not Null	Constraint	Comment
COMMODITY_ID	Identifier for the commodity	INTEGER	N		FK – COMMODITY_ LUT
COMMODITY_CATEGORY_ID	Identifier for the commodity category	INTEGER	Y		FK – COMMODITY_CATEGORY_LUT
SHELF_LIFE_RANGE_DAYS	Shelf life range for the commodity/commodity category in days	TEXT(255)	Y		
SHELF_LIFE_DAYS	Shelf life in days for the commodity/commodity category (currently using maximum of range)	NUMBER	Y		

SUSCEPTIBLE_POPULATION

Contains susceptible data. (Pathogen Category)

Field Name	Description	Data Type	Not Null	Constraint	Comment
PATHOGEN_CATEGORY_ID	Identifier for the pathogen category	INTEGER	Y	PK	FK – PATHOGEN_CATEGORY_LUT
DEFAULT_BIN	Default bin value	TEXT(50)	Y		
USER_BIN	User defined bin value, default value is DEFAULT_BIN, can be manually changed by the user in this table	TEXT(50)	Y		
USER_BIN_RANK	Rank assigned to the pathogen category combination for the susceptible population based on where the value in USER_BIN falls in the user defined bins	INTEGER	Y		