



City of Albuquerque Information Technology Services Division Data Management

Data.cabq.gov Core Metadata Requirements

Contact Information

Who is the contact for this dataset? The contact will be the City employee who is accountable for the data provided in this dataset and can act as front-line support in the event of any questions about the data.

Name	Lorie Stoller
Department/Division	EHD / Consumer Health
Phone	768-2718
Email	lstoller@cabq.gov

What Does this Dataset Describe?

What is the name of this dataset? How should a user identify this dataset in any communication with contact above? Provide a shorter description of the Dataset that can act as a one-line summary of the dataset when dealing with stakeholders. Provide a longer description of the data that can be readily understood by non-technical users.

Dataset Title	LIVES
Short Description	Restaurant Inspection Results
Full Non-Technical Description	
Restaurant Inspection Results in the YELP Local Inspector Value-Entry Specification (LIVES) format. Please see http://www.yelp.com/healthscores for more details on the LIVES format.	

How Should this Dataset be Cited?

How should external sources refer to this dataset in publications or documentation? Often this will simply be the URL and the date retrieved.

http://data.cabq.gov/business/LIVES/coa_LIVES.Zip

businesses.csv
inspections.csv
violations.csv
feed_info.csv
legend.csv

Does the Dataset Reflect a Particular Time Period?

Provide any date restrictions that may affect the validity of the data. The table fields are defined as follows:

<i>Field</i>	<i>Definition</i>
<i>Start Date</i>	<i>Start date of the time period within which this data falls. Format: MM/DD/YYYY HH:MM:SS.</i>
<i>End Date</i>	<i>End date of the time period within which this data falls. Format: MM/DD/YYYY HH:MM:SS.</i>
<i>Dataset Refresh Interval</i>	<i>Time period between Dataset refreshes. Format: “nn [seconds/minutes/hours/days/weeks/months/years]” or the word “Static” if never refreshed.</i>
<i>Data Expiration Date</i>	<i>Date after which the data must be considered stale and no longer of sufficient utility (fit-for-purpose). Format: MM/DD/YYYY HH:MM:SS.</i>
<i>Dataset Review Date</i>	<i>Date after which this dataset will be reviewed by the City for utility (fit-for-purpose) and usage. Format: MM/DD/YYYY HH:MM:SS.</i>
<i>Comments</i>	<i>Specific comments related to any time-specific features of this dataset.</i>

Start Date	System date – 1 year
End Date	current
Dataset Refresh Interval	Nightly - ACTIVITY_DATE < SYSDATE - 14
Dataset Expiration Date	
Dataset Review Date	Production date
Comments	
Inactive facilities are removed – based on status – out of business on ASP application	

Dataset Definition/Format

Provide a field-by-field breakdown and definition of each record. This section acts as the formal data dictionary for an individual record.

Field Name	Format	Description
businesses.csv		
business_id	xs:string length = 20	FACILITY_ID > ID of the facility where the inspection occurred
name	xs:string length = 202	FACILITY_NAME > Name of the facility where the inspection occurred
address	xs:string length = 102	SITE_ADDRESS > Street address of the facility where the inspection occurred
city	xs:string length = 52	CITY > City of the facility where the inspection occurred
state	xs:string length = 6	STATE > State of the facility where the inspection occurred
postal_code	xs:string length = 22	ZIP > Zip of the facility where the inspection occurred
phone_number	xs:string length = 42	PHONE > Phone number of the facility where the inspection occurred
inspections.csv		
business_id	xs:string length = 20	FACILITY_ID > ID of the facility where the inspection occurred
score	string	Based on ACTION_CODE when '09' then '100' when '11' then '100' else '0' This field is NOT to be displayed on any application, report, etc. If you want to display information related to the outcome of the inspection you MUST use the following translation: 0 = Unsatisfactory; 100 = Approved. The City of Albuquerque has a requirement to not display a score for the inspections. This field is to be used in conjunction with the legend.csv file.
date	dateTime	INSPECTION_DATE > Date the inspection occurred - YYYYMMDD
description	string	If Violation Count is null then ACTION_DESCR If Violation Count is 0 then ACTION_DESCR Otherwise blank
type	string	Based on INSPECTION_DESC <ul style="list-style-type: none"> 'New Business Inspection' then 'routine' 'ROUTINE FOOD INSPECTION' then 'routine' 'FOOD FOLLOW-UP INSPECTION' then 'follow-up' 'PRE-OPENING FOOD INSPECTION' then 'initial' If not one of the above then INSPECTION_DESC

violations.csv		
business_id	xs:string length = 20	FACILITY_ID > ID of the facility where the inspection occurred
date	dateTime	INSPECTION_DATE > Date the inspection occurred – YYYYMMDD
code	String	VIOLATION_CODE > The code used by the City of Albuquerque to indicate the violation that was found.
description	xs:string length = 202	MEMO > Description of the violation that was found
feed_info.csv		
feed_date	date	System Date – updated each time a dataset is created
feed_version	string	1.0 – Version of the LIVES used
municipality_name	string	Albuquerque
municipality_url	string	www.cabq.gov
contact_email	string	ABQ-Data@cabq.gov
legend.csv		
minimum_score	number	Minimum score that can be classified with this description. Available values: 0, 100
maximum_score	number	Maximum score that can be classified with this description. Available values: 0, 100
description	string	Formatted version of the score in the format typically presented by the municipality. Available values: 0 = Unsatisfactory ; 100 = Approved

Dataset Technical Description

Provide a technical description of the dataset. This should be a complete technical description aimed at developers and expert users who need to understand the scope, strengths and limitations of the dataset.

Csv format with header row for field names

UTF-16 format

Filters Used:

All files

- CURRENT_STATUS = '01'
- PE not in ('0701', '0702', '0703', '0704')
- Inspection_type is null or not equal 9

- Service not in ('901', '902', '909')
- Service in ('100', '101', '109', '110')

businesses.csv – substring(PROGRAM_CATEGORY,1,2) in ('04','05','06','10')

inspections.csv – substring(PROGRAM_CATEGORY,1,2) in ('04','05','06','10')

– ACTION_CODE in ('06', '09', '10', '11', '17')

– VIOLATION_CODE not in ('59', '04 59', '55', '68', '56', '04 S57', '09 I1')

– ACTIVITY_DATE > 12 Months from extract date

violations.csv – substring(PROGRAM_CATEGORY,1,2) in ('04','05','06','10')

– violation_code not in ('04 00', ' 67 ', ' 00 ')

– ACTION_CODE in ('06', '09', '10', '11', '17')

– violation_code not in ('59', '04 59', '55', '68', '56', '04 S57', '09 I1')

– ACTIVITY_DATE > 12 Months from extract date

Dataset Assumptions

What technical and business assumptions are implied in the creation of this dataset? Examples could include the way in which a salary figure was calculated or data that was omitted for a specific reason.

None -

Who Produced the Dataset?

Which department in the City produced this dataset? Note that this might not always be the data owner. An example of this could be a dataset that ITSD produced on behalf of EHD who owned the data.

ITSD – Business Intelligence Group

Who Manages the Data?

Where did this data originate? Who manages the data used in this dataset? Note that this might not always be the dataset producer. An example of this could be a dataset that ITSD produced on behalf of EHD who manages the data.

–Environmental Health Department

Why was the Dataset Created?

All datasets should have an explicit reason for existence and should, somehow, have value to someone.

What is the perceived value that this dataset will bring?

EHD is charged with protecting public health and tracking inspection results. The data provides information to citizens on the pass / fail status of establishments.

How was the Dataset Created?

How was this dataset produced? Was it a manual process? An automated process? What were the main IT systems involved in producing this dataset?

Data is generated from the production EHD application and sent to BI group then extracted into csv

What Similar or Related Data Should the User be Aware of?

Are there any other datasets available that may contain related or similar information? Might there be situations in which these other datasets might be a better alternative?

bizinfo.cabq.gov

<http://data.cabq.gov/business/foodinspections/FoodInspectionsCABQ-en-us.csv>

<http://data.cabq.gov/business/foodinspections/FoodInspectionsCABQ-en-us.xml>

How Reliable are the Data?

Are there any concerns about overall data reliability? Are there any data problems that the user needs to be aware of? Are there any constraints with data accuracy? What levels of confidence with this dataset could the user reasonably assume?

Data is produced from the production EHD application that is used on a daily basis.

How Well Have the Observations Been Checked?

What quality assurance steps have been performed? Sometimes, a third-party verification/audit process may also be required. If so, provide the name of the third-party who performed the verification.

Data has been verified by the department – Inspectors enter inspections daily, then the supervisors review the inspections for spelling and grammar errors, and verify the codes used. This process take up to 14 days so records are included in the extract where ACTIVITY_DATE < SYSDATE – 14.

Are there Legal Restrictions on the Access or Use of the Data?

Are there any specific legal or compliance restrictions for this data? How might this affect the way in which end users might access and use this data?

Hippa data is not included.

Legal Disclaimer

The City's standard copyright, disclaimers and legal statements may be found at

<http://www.cabq.gov/about/legal>. The City data policy governing data.cabq.gov may be found at <http://data.cabq.gov/policy/>.