

# Food\_Inspections\_CA

## PART 1:

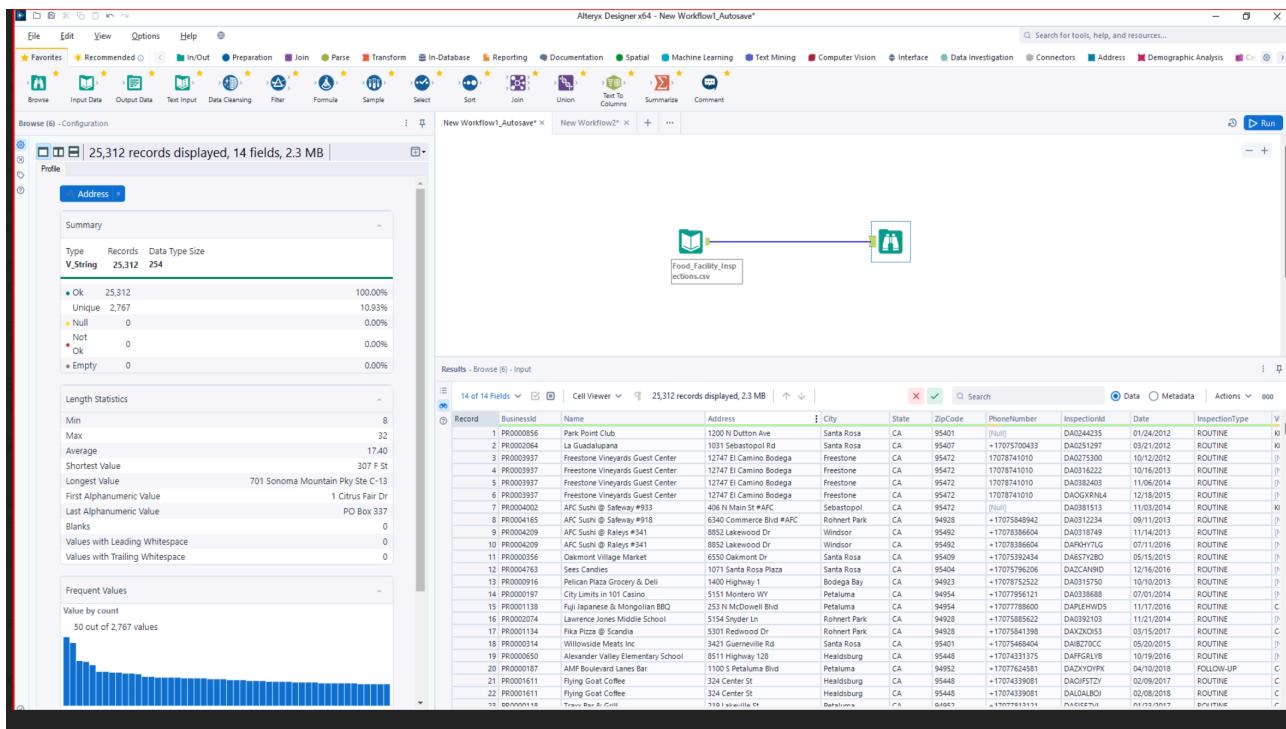
### Overview

#### Dataset statistics

Number of variables	13
Number of observations	25312
Missing cells	24530
Missing cells (%)	7.5%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	2.5 MiB
Average record size in memory	104.0 B

#### Variable types

Text	8
Categorical	3
Numeric	1
DateTime	1

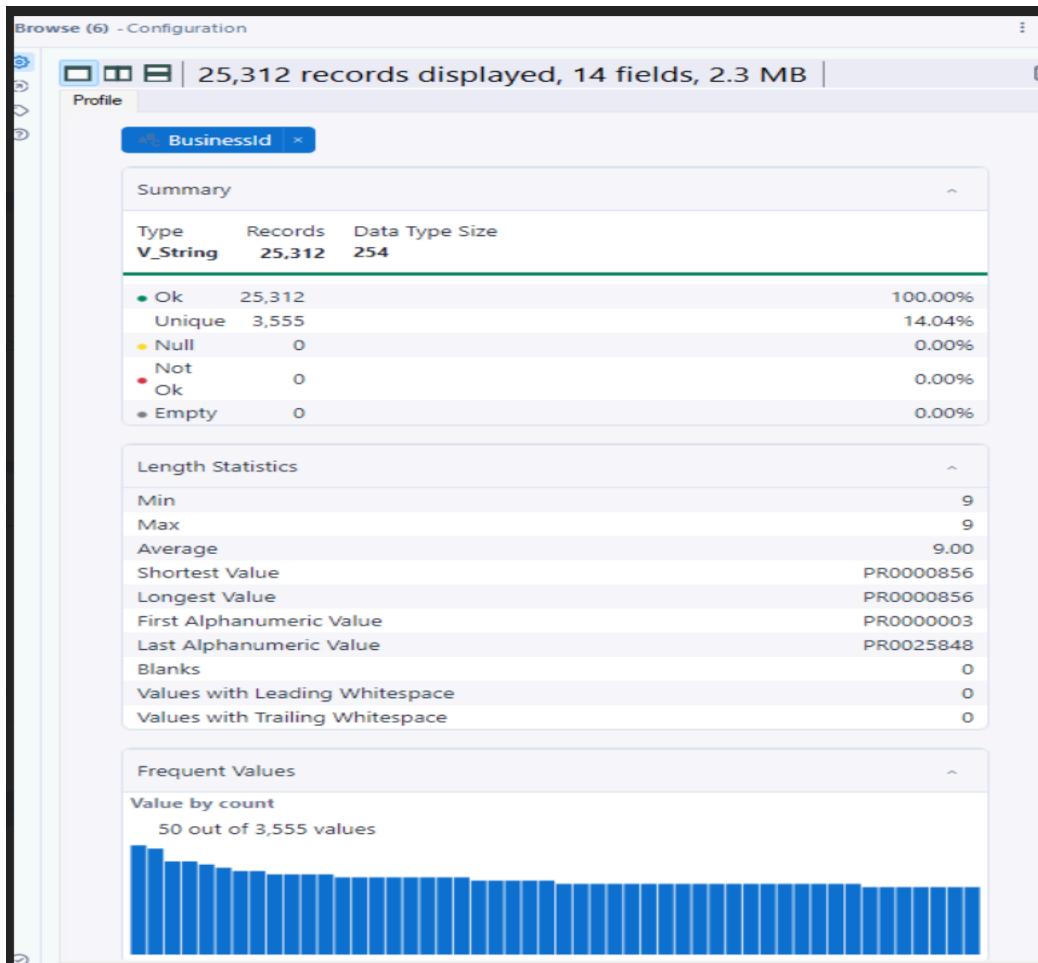


## Variables:

### BusinessId

Text

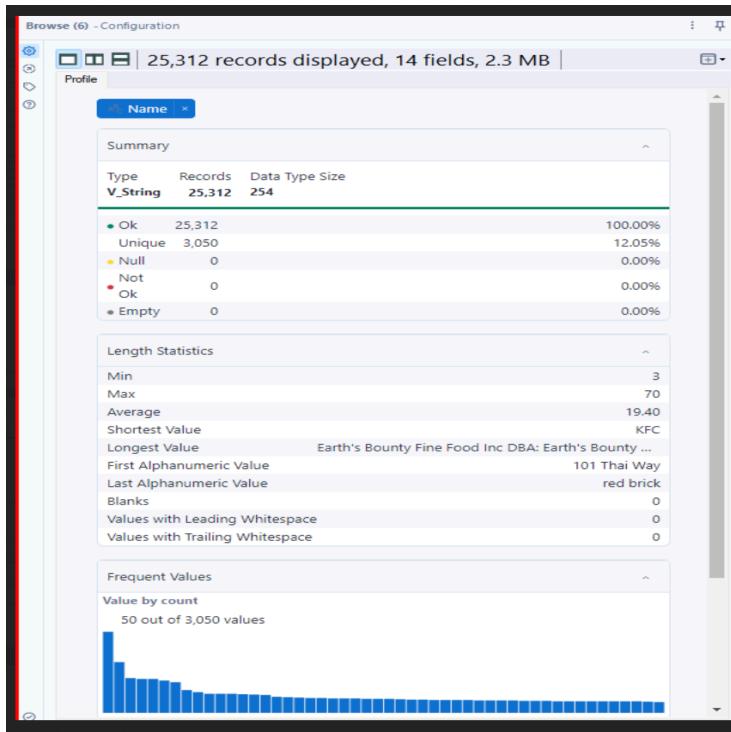
Distinct	3555
Distinct (%)	14.0%
Missing	0
Missing (%)	0.0%
Memory size	197.9 KiB



## Name

Text

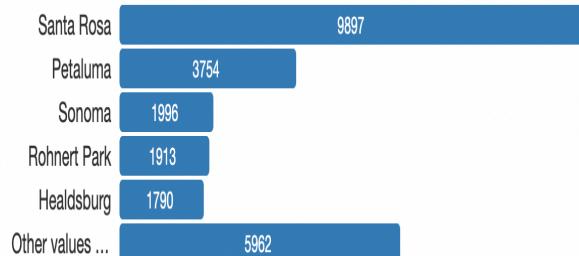
Distinct	3050
Distinct (%)	12.0%
Missing	0
Missing (%)	0.0%
Memory size	197.9 KiB



## City

Categorical

Distinct	43
Distinct (%)	0.2%
Missing	0
Missing (%)	0.0%
Memory size	197.9 KIB



More details

Profile

City

Summary

Type	Records	Data Type	Size
V_String	25,312	254	

Ok	25,312	100.00%
Unique	43	0.17%
Null	0	0.00%
Not Ok	0	0.00%
Empty	0	0.00%

Length Statistics

Min	4
Max	17
Average	9.30
Shortest Value	Napa
Longest Value	Boyes Hot Springs
First Alphanumeric Value	Agua Caliente
Last Alphanumeric Value	Windsor
Blanks	0
Values with Leading Whitespace	0
Values with Trailing Whitespace	0

Frequent Values

Value by count  
43 out of 43 values



## State

Categorical

HIGH CORRELATION IMBALANCE

Distinct	2	CA	25310
Distinct (%)	<0.1%		^ 2
Missing	0		
Missing (%)	0.0%		
Memory size	197.9 KIB		

□ □ □ | 25,312 records displayed, 14 fields, 2.3 MB | E

Profile State

Summary

Type	Records	Data Type	Size
V_String	25,312	254	

• Ok	25,312	100.00%
Unique	2	0.01%
• Null	0	0.00%
Not Ok	0	0.00%
• Empty	0	0.00%

Length Statistics

Min	1
Max	2
Average	2.00
Shortest Value	
Longest Value	CA
First Alphanumeric Value	CA
Last Alphanumeric Value	
Blanks	0
Values with Leading Whitespace	0
Values with Trailing Whitespace	0

Frequent Values

Value by count

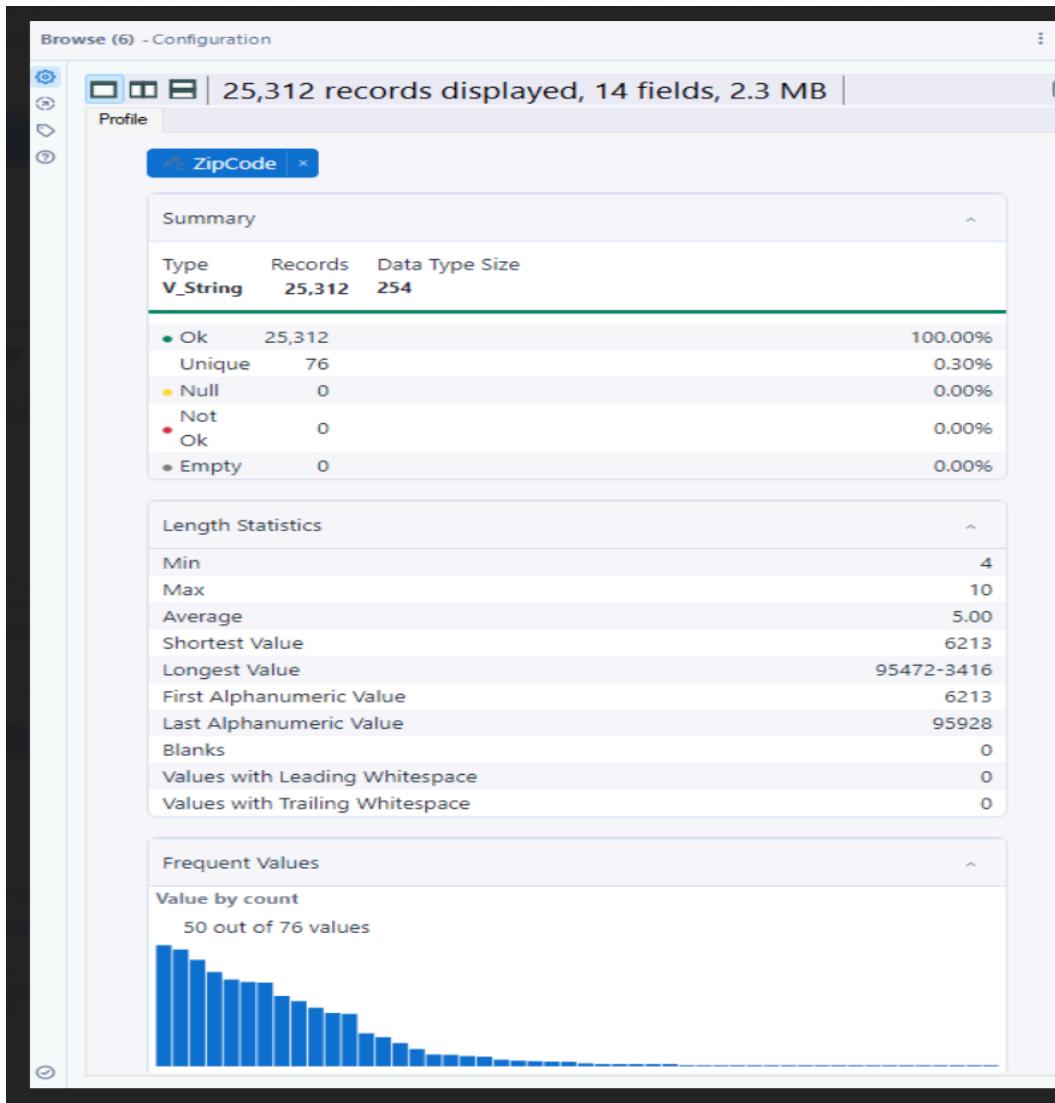
2 out of 2 values



## ZipCode

Text

Distinct	76
Distinct (%)	0.3%
Missing	0
Missing (%)	0.0%
Memory size	197.9 KIB



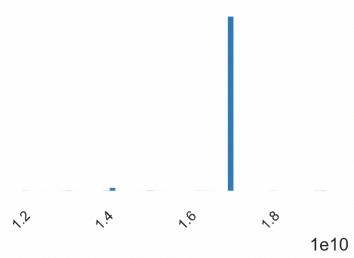
## PhoneNumber

Real number (R)

HIGH CORRELATION MISSING

Distinct	2556
Distinct (%)	12.2%
Missing	4307
Missing (%)	17.0%
Infinite	0
Infinite (%)	0.0%
Mean	$1.698409 \times 10^{10}$

Minimum	$1.2019703 \times 10^{10}$
Maximum	$1.9499439 \times 10^{10}$
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	197.9 KiB



25,312 records displayed, 14 fields, 2.3 MB

Profile PhoneNumber

Summary

Type	Records	Data Type Size
V_String	25,312	254

Ok	21,005	82.98%
Unique	3,353	13.25%
Null	4,307	17.02%
Not Ok	0	0.00%
Empty	0	0.00%

Length Statistics

Min	11
Max	12
Average	11.80
Shortest Value	17078741010
Longest Value	+17075700433
First Alphanumeric Value	+12019702799
Last Alphanumeric Value	19258907977
Blanks	0
Values with Leading Whitespace	0
Values with Trailing Whitespace	0

Frequent Values

Value by count

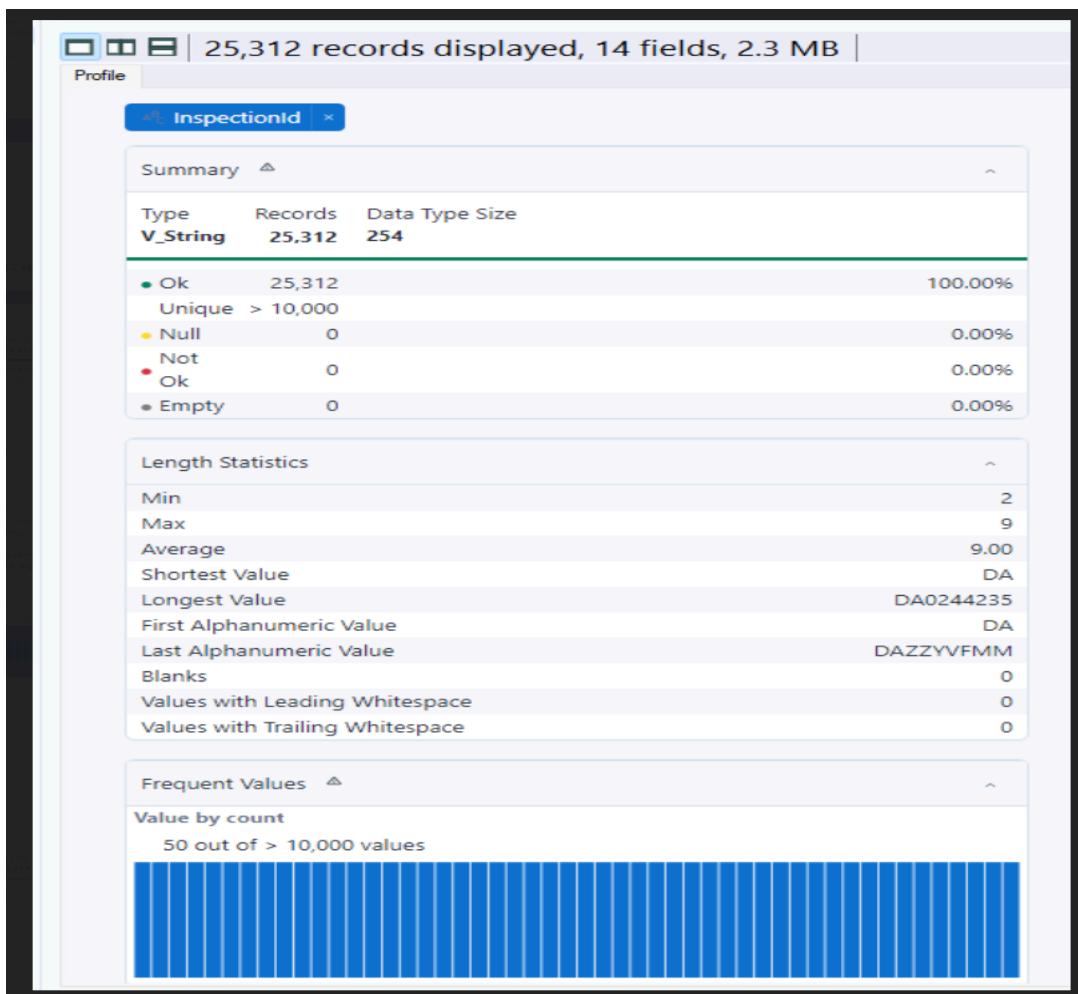
50 out of 3,353 values

## InspectionId

Text

UNIQUE

Distinct	25312
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%
Memory size	197.9 KiB

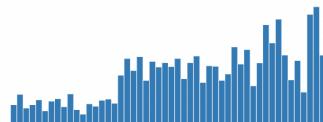


## Date

Date

Distinct	2219
Distinct (%)	8.8%
Missing	3
Missing (%)	< 0.1%
Memory size	197.9 KiB

Minimum	2009-07-02 00:00:00
Maximum	2018-07-02 00:00:00



2011-05-13 03:06:40

2014-05-13 22:53:20

2017-07-13 22:40:00

**Date**

**Summary**

Type	Records	Data Type	Size
V_String	25,312	254	

Ok	25,309	99.99%
Unique	2,219	8.77%
Null	3	0.01%
Not Ok	0	0.00%
Empty	0	0.00%

**Length Statistics**

Min	10
Max	10
Average	10.00
Shortest Value	01/24/2012
Longest Value	01/24/2012
First Alphanumeric Value	01/02/2013
Last Alphanumeric Value	12/31/2015
Blanks	0
Values with Leading Whitespace	0
Values with Trailing Whitespace	0

**Frequent Values**

Value by count  
50 out of 2,219 values

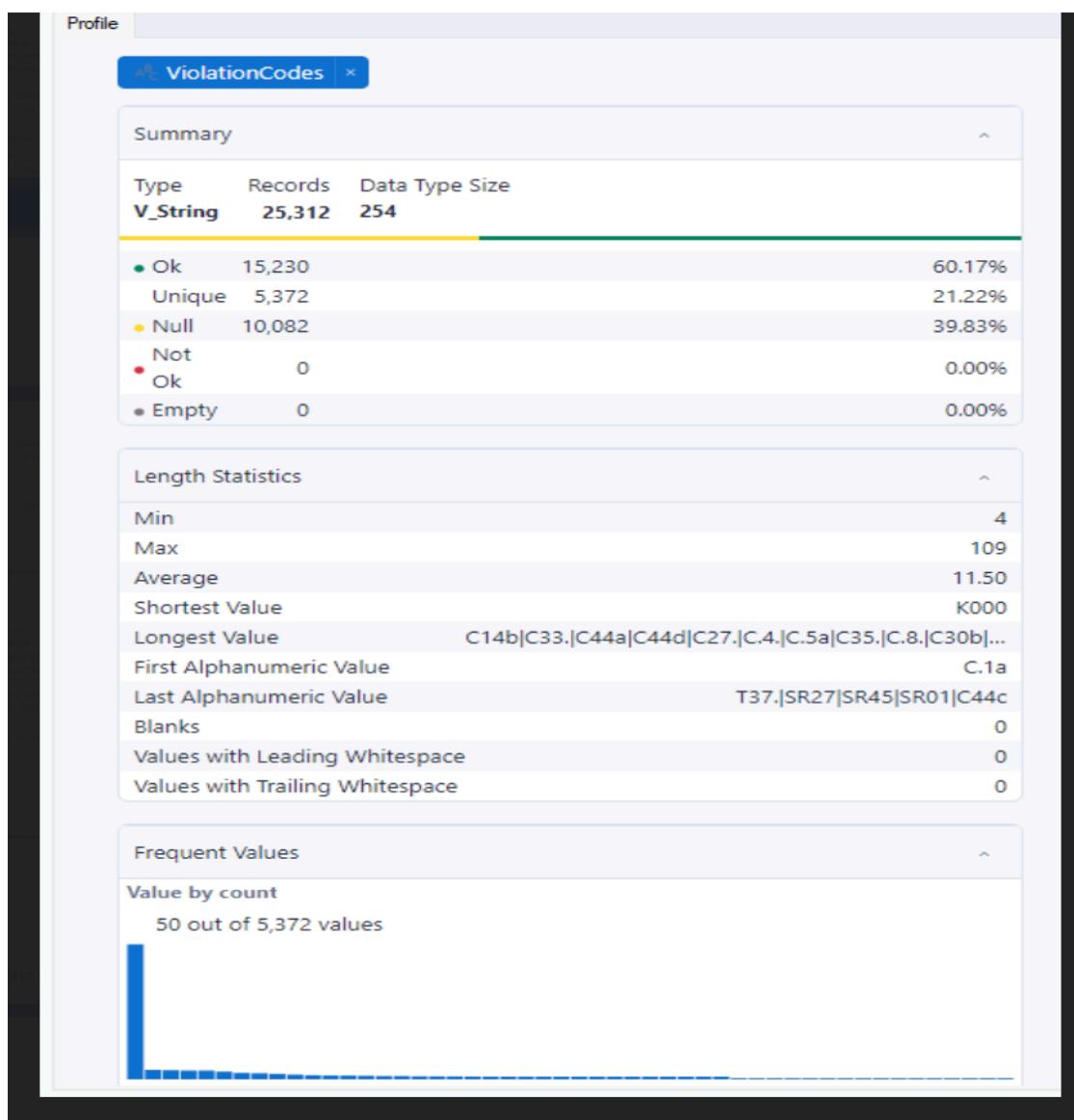
A bar chart showing the top 50 frequent values for Date. The x-axis lists the values, and the y-axis shows their frequency. The values are mostly dates, with one notable outlier being a string value.

## ViolationCodes

Text

MISSING

Distinct	5372
Distinct (%)	35.3%
Missing	10082
Missing (%)	39.8%
Memory size	197.9 KiB

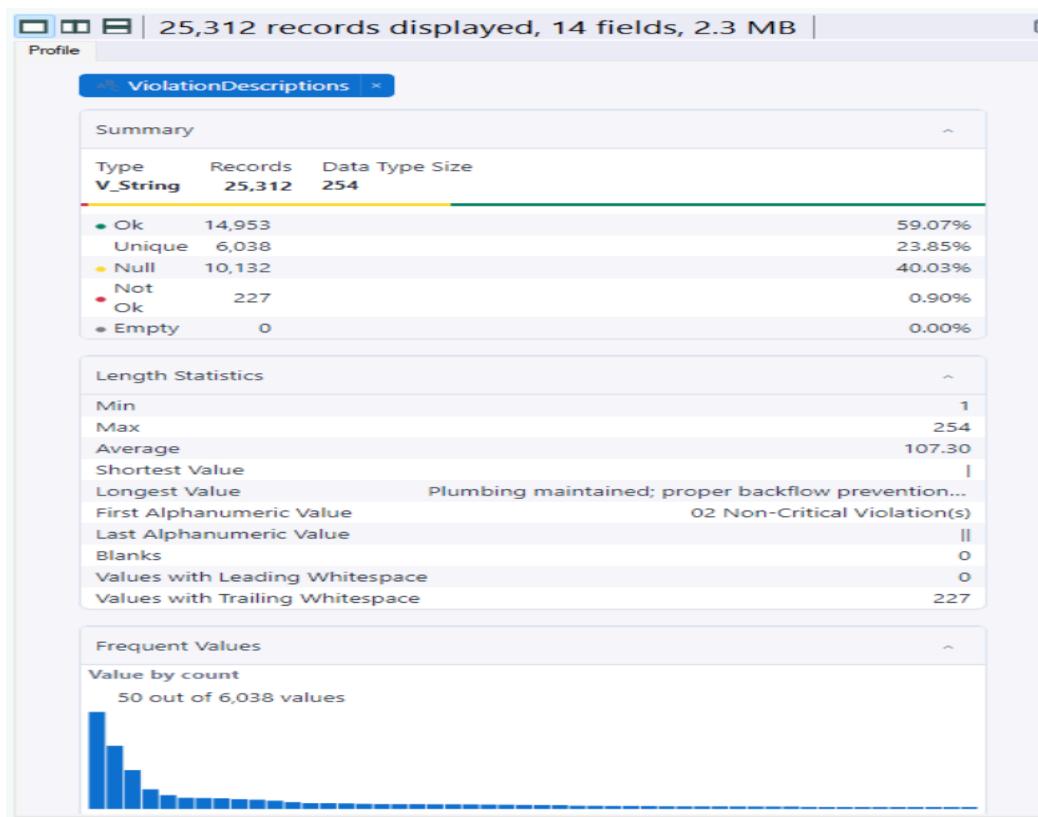


## ViolationDescriptions

Text

MISSING

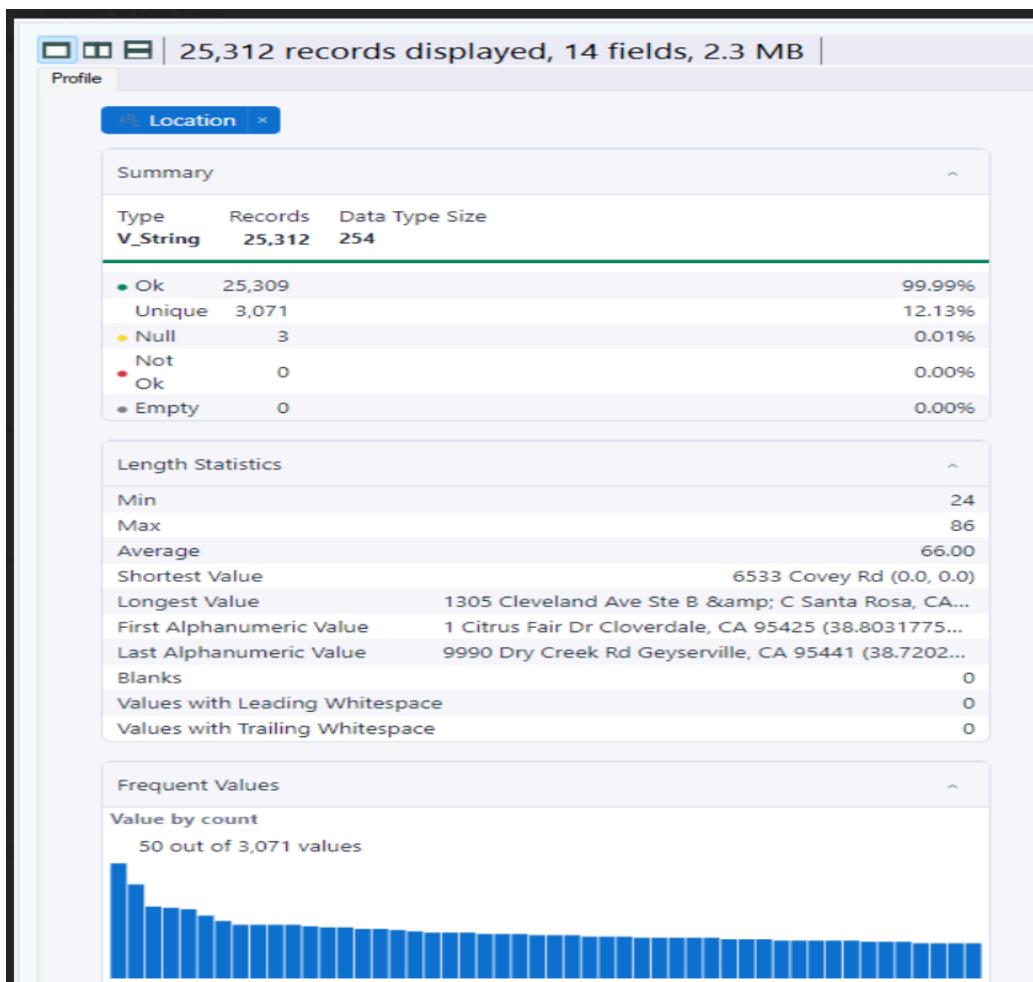
Distinct	6039
Distinct (%)	39.8%
Missing	10132
Missing (%)	40.0%
Memory size	197.9 KiB



## Location

Text

Distinct	3071
Distinct (%)	12.1%
Missing	3
Missing (%)	< 0.1%
Memory size	197.9 KiB



## InspectionType

Categorical

### IMBALANCE

Distinct	3
Distinct (%)	< 0.1%
Missing	3
Missing (%)	< 0.1%
Memory size	197.9 kB



### InspectionType

#### Summary

Type	Records	Data Type	Size
V_String	25,312	254	

● Ok	25,309	99.99%
● Unique	3	0.01%
● Null	3	0.01%
● Not	0	0.00%
● Ok	0	0.00%
● Empty	0	0.00%

#### Length Statistics

Min	7
Max	9
Average	7.20
Shortest Value	ROUTINE
Longest Value	FOLLOW-UP
First Alphanumeric Value	COMPLAINT
Last Alphanumeric Value	ROUTINE
Blanks	0
Values with Leading Whitespace	0
Values with Trailing Whitespace	0

#### Frequent Values

Value by count  
3 out of 3 values

A horizontal bar chart showing the frequency of inspection types. The first bar is large and blue, the second is smaller and blue, and the third is very small and light blue.

## MISSING VALUES

