Alteryx Workflow – Midterm Team Project – Food Inspection Facility Data

A computer network diagram with a cable connected to a computer

Description automatically generated with medium confidence

* Let’s Describe data inconsistencies and missing data values -
* and How we tackled the problems with Alteryx Tools -

1. Businessid - Unique Identifier for Health services of a business. It can be concluded that the ID PR0021173 is majorly occurring.

Data type - text - string

Min and max - 9

No null values

A screenshot of a graph

Description automatically generated

1. Name - Food Facility Name – There’s about 175,864 records that are populated in the column name.

Data type - Text - string

Min - 3

Max – 70

A screenshot of a computer

Description automatically generated

1. Address – Predominantly the Address 113 Healdsburg is seen to have the most inspections for food facilities.

Data type - Text - string

Min - 8

Max – 32

A screenshot of a graph

Description automatically generated

1. City of business – Santa Rosa is the city where the FoodFacility inspections have majorly took place

Data type - Text - string

Min - 4

Max – 17

A screenshot of a computer screen

Description automatically generated

1. State - state of Business – Predominantly in CA - California

Data type - Text - string

Min – 1

Max - 2

1. Zipcode - Zipcode where the healthcare business runs.

Data type - Text - string

Min - 4

Max - 10

A screenshot of a table

Description automatically generated

1. PhoneNumber – With about 24,876 Null – Empty cells, With data cleasing tool we can fix nulls to blank string/int values

Data type - Text - string

Min - 11

Max - 12

A screenshot of a computer

Description automatically generated

1. InspectionId - Unique code to identify inspection services -

Data type - Text - string

Min - 2

Max - 9

1. Date - Date of Inspection – September 1st of 2018 is a day with about 1000+ Inspection cases.

Floating time stamp

Min - 10

Max - 10

3 null values – overcame by Data cleanse tool

A screenshot of a graph

Description automatically generated

1. Inspection type - Type of inspection - categorised to routine, follow up and complaint and its deduced that Routine checks are common followed by inspection follow-ups

Data type - text. - string

Min - 7

Max - 9

3 null values – data cleanse

A screenshot of a computer

Description automatically generated

1. Violation code - Code used by health services to identify the violation

Data type - text. - string

Min - 4

Max - 4

10,082 null values – Used Data cleanse tool to fill null values with blank strings.

1. Violation description - categorised to critical, non-critical ,walls , ceilings and floor damages.

Data type - text. - string

Min - 1

Max - 100

10,194 null values – Null values are removed by blank strings

A screenshot of a data

Description automatically generated

1. Location - Place where inspection took place for food facilities.

Data type - text. - string

Min - 24

Max - 86

3 null values – Data cleanse tool used.

1. DI\_CreatedDate – DateTime when row is loaded

Data type - Date

Min - 100

Max - 100

1. DI\_WorkflowFileName – Filename that you use to load the data

Data type - String

Min - 100

Max - 100

16. DI\_Workflow\_ProcessID – Workflow / Job ID

Data type - String

Min - 100

Max - 100