

San Francisco Trees Data Profiling & Data Prep

The screenshot shows the Alteryx Designer x64 interface with a new workflow. The 'Input Data (1) - Configuration' pane on the left shows the connection to 'C:\Users\moira\Downloads\Street_Tree_List (1).txt'. The 'Resolve File Type' dialog box is open, asking to resolve the file type. The file is not recognized, and the user is prompted to choose a type. The 'Read it as a delimited text file' option is selected, with 'Comma' as the delimiter and 'First Row Contains Field Names' checked. The 'Results - Input Data (1) - Output' pane shows 0 of 0 fields.

Alteryx Designer x64 - New Workflow*

File Edit View Options Help

Search for tools, help, and resources...

Input Data (1) - Configuration

Connect a File or Database

C:\Users\moira\Downloads\Street_Tree_List (1).txt

Set Up a Connection

Use Data Connection Manager (DCM)

Options

Name	Value
1 Record Limit	
2 File Format	Alteryx Database (*.yxd)
3 Search SubDirs	<input type="checkbox"/>
4 Output File Name as Field	No

Preview (first 100 records)

Refresh

Resolve File Type

The selected file is not a recognized type

C:\Users\moira\Downloads\Street_Tree_List (1).txt

Read it as a built-in type

Alteryx Database (*.yxd)

Read it as a fixed width text file

Read it as a delimited text file

Delimiter

Comma ☒ Tab ☐ Space ☐ None

Other:

First Row Contains Field Names ☒

OK Cancel Help

Results - Input Data (1) - Output

0 of 0 Fields

Cell Viewer

No data available. Use Ctrl+R to run the workflow.

The screenshot shows the Alteryx Designer x64 interface with a new workflow. The 'Browse (4) - Configuration' pane on the left shows the profile of the data, including 196,766 records displayed, 24 fields, 23 MB. The 'Results - Browse (4) - Input' pane shows a table with 24 of 24 fields and 196,766 records displayed, 23 MB. The table has columns: Record, TreeID, qLegalStatus, qSpecies, qAddress, SiteOrder, and qSiteInfo. The workflow diagram shows a sequence of tools: 'Street_Tree_List (1).txt', 'Filter', 'Join', 'Text To Columns', and 'Summarize'.

Alteryx Designer x64 - New Workflow*

File Edit View Options Help

Search for tools, help, and resources...

Browse (4) - Configuration

196,766 records displayed, 24 fields, 23 MB

Profile

196,766 records displayed, 24 fields, 23 MB

TreeID

1	1
2	1
3	1
11	1
21	1

995 more >

qLegalStatus

DPW Maintained	143,396
Permitted Site	40,133
Undocumented	8,619
Significant Tree	1,901
Planning Code 138.1 required	1,123

7 more >

qSpecies

Trees(s) :	11,811
Platanus x hispanica : Sycamore: London Plane	11,798
Lophostemon confertus : Brisbane Box	8,927

New Workflow*

Street_Tree_List (1).txt

Results - Browse (4) - Input

24 of 24 Fields

196,766 records displayed, 23 MB

Record	TreeID	qLegalStatus	qSpecies	qAddress	SiteOrder	qSiteInfo
1	168225	DPW Maintained	Arbutus 'Marina' : Hybrid Strawberry Tree	2547 Vallejo St	1	Sidewalk: Curb side : Cut
2	168228	DPW Maintained	Arbutus 'Marina' : Hybrid Strawberry Tree	2547 Vallejo St	4	Sidewalk: Curb side : Cut
3	189160	DPW Maintained	Alrocarpus gracilior : Fern Pine	100 Kissing St	3	:
4	189158	DPW Maintained	Alrocarpus gracilior : Fern Pine	100 Kissing St	1	:
5	189161	DPW Maintained	Alrocarpus gracilior : Fern Pine	100 Kissing St	4	:
6	168226	DPW Maintained	Arbutus 'Marina' : Hybrid Strawberry Tree	2547 Vallejo St	2	Sidewalk: Curb side : Cut
7	269465	Undocumented	Thuja occidentalis 'Emerald' : Emerald Arborvitae	1670 09th Ave	1	Sidewalk: Curb side : Cut
8	269478	Planning Code 138.1 required	Magnolia grandiflora 'Little Gem' : Little Gem Ma...	99h Tandang Sora	1	Sidewalk: Curb side : Cut
9	269480	Planning Code 138.1 required	Magnolia grandiflora 'Little Gem' : Little Gem Ma...	99h Tandang Sora	3	Sidewalk: Curb side : Cut
10	253378	Planning Code 138.1 required	Platanus x hispanica : Sycamore: London Plane	2 Alpine Ter	2	Sidewalk: Curb side : Cut
11	253377	Planning Code 138.1 required	Platanus x hispanica : Sycamore: London Plane	2 Alpine Ter	1	Sidewalk: Curb side : Cut
12	242757	DPW Maintained	Ulmus parvifolia : Chinese Elm	194 Corbett Ave	1	Sidewalk: Curb side : Cut
13	191067	DPW Maintained	Eriobotrya deflexa : Bronze Loquat	374 05th St	3	Sidewalk: Curb side : Cut
14	245047	DPW Maintained	Leptospermum laevigatum : Australian Tea Tree	1034 Page St	2	Sidewalk: Curb side : Cut
15	168227	DPW Maintained	Arbutus 'Marina' : Hybrid Strawberry Tree	2547 Vallejo St	3	Sidewalk: Curb side : Cut

Alteryx Designer x64 - New Workflow*

Search for tools, help, and resources...

File Edit View Options Help

Favorites Recommended In/Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Browse Input Data Output Data Text Input Data Cleansing Filter Formula Sample Select Sort Join Union Text To Columns Summarize Comment

Formula (5) - Configuration

Output Column Data Preview

1 DI_Create_Date 2023-02-07

DateTimeNow()

Data type: Date Size: 10

New Workflow* x

Street_Tree_List (1).tsv

Results - Formula (5) - Input

24 of 24 Fields 4,113 of 196,766 records displayed (partial results)

Record	TreeID	qLegalStatus	qSpecies	qAddress	SiteOrder	qSiteInfo
1	168225	DPW Maintained	Arbutus 'Marina' : Hybrid Strawberry Tree	2547 Vallejo St	1	Sidewalk: Curb side : Cut
2	168228	DPW Maintained	Arbutus 'Marina' : Hybrid Strawberry Tree	2547 Vallejo St	4	Sidewalk: Curb side : Cut
3	189160	DPW Maintained	Alrocarpus gracilior : fern Pine	100 Kissing St	3	:
4	189158	DPW Maintained	Alrocarpus gracilior : fern Pine	100 Kissing St	1	:
5	189161	DPW Maintained	Alrocarpus gracilior : fern Pine	100 Kissing St	4	:
6	168226	DPW Maintained	Arbutus 'Marina' : Hybrid Strawberry Tree	2547 Vallejo St	2	Sidewalk: Curb side : Cut
7	269465	Undocumented	Thuja occidentalis 'Emerald' : Emerald Arborvitae	1670 09th Ave	1	Sidewalk: Curb side : Cut
8	269478	Planning Code 138.1 required	Magnolia grandiflora 'Little Gem' : Little Gem Ma...	99x Tandang Sora	1	Sidewalk: Curb side : Cut
9	269480	Planning Code 138.1 required	Magnolia grandiflora 'Little Gem' : Little Gem Ma...	99x Tandang Sora	3	Sidewalk: Curb side : Cut
10	253378	Planning Code 138.1 required	Platanus x hispanica : Sycamore: London Plane	2 Alpine Ter	2	Sidewalk: Curb side : Cut
11	253377	Planning Code 138.1 required	Platanus x hispanica : Sycamore: London Plane	2 Alpine Ter	1	Sidewalk: Curb side : Cut
12	242757	DPW Maintained	Ulmus parvifolia : Chinese Elm	194 Corbett Ave	1	Sidewalk: Curb side : Cut
13	191067	DPW Maintained	Eriobotrya deflexa : Bronze Loquat	374 05th St	3	Sidewalk: Curb side : Cut
14	245047	DPW Maintained	Leptospermum laevigatum : Australian Tea Tree	1034 Page St	2	Sidewalk: Curb side : Cut
15	168227	DPW Maintained	Arbutus 'Marina' : Hybrid Strawberry Tree	2547 Vallejo St	3	Sidewalk: Curb side : Cut

2°C Mostly clear

Search

3:00 PM 2/7/2023

Alteryx Designer x64 - New Workflow*

Search for tools, help, and resources...

File Edit View Options Help

Favorites Recommended In/Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Browse Input Data Output Data Text Input Data Cleansing Filter Formula Sample Select Sort Join Union Text To Columns Summarize Comment

Browse (6) - Configuration

196,766 records displayed, 25 fields, 23 MB

Profile

196,766 records displayed, 25 fields, 23 MB

TreeID

1 1

2 1

3 1

11 1

21 1

995 more >

qLegalStatus

DPW Maintained 143,396

Permitted Site 40,133

Undocumented 8,619

Significant Tree 1,901

Planning Code 138.1 required 1,123

7 more >

qSpecies

Tree(s) : 11,811

Platanus x hispanica : Sycamore: Lon... 11,798

Lophostemon confertus : Brisbane B... 8,927

New Workflow* x

Street_Tree_List (1).tsv

DI_Create_Date = DateTimeNow()

Results - Browse (6) - Input

25 of 25 Fields Cell Viewer 196,766 records displayed, 23 MB

Record	Fire Prevention Districts	Police Districts	Supervisor Districts	Zip Codes	Neighborhoods (old)	Analysis Neighborhoods	DI_Create_Date
1 9930563	13	9	1	57	27	30	2023-02-07
2 133225	13	9	1	57	27	30	2023-02-07
3 744545	8	2	9	28853	19	20	2023-02-07
4 332298	8	2	9	28853	19	20	2023-02-07
5 3345712	8	2	9	28853	19	20	2023-02-07
6 1974892	13	9	1	57	27	30	2023-02-07
7 183035	1	8	4	56	12	14	2023-02-07
8 [Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	2023-02-07
9 [Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	2023-02-07
10 122497	15	5	5	29492	9	3	2023-02-07
11 122497	15	5	5	29492	9	3	2023-02-07
12 16181	15	5	5	28862	3	5	2023-02-07
13 117102	14	2	9	28856	34	34	2023-02-07
14 152145	15	5	11	29492	9	3	2023-02-07
15 13267	13	9	1	57	27	30	2023-02-07

2°C Mostly clear

Search

3:02 PM 2/7/2023

Alteryx Designer x64 - New Workflow*

Formula (7) - Configuration

Output Column: DI_Job_Name [Engine.WorkflowFileName]

Data type: V_String Size: 1073741823

Workflow Diagram: Street_Tree_List (1).tsv -> Join -> Formula (7) -> Output Column: DI_Job_Name

Results - Formula (7) - Input

Record	qLegalStatus	qSpecies	qAddress	SiteOrder	qSiteInfo	PlantType	qCaretai
1	DPW Maintained	Arbutus 'Marina' : Hybrid Strawberry Tree	2547 Vallejo St	1	Sidewalk: Curb side : Cutout	Tree	Private
2	DPW Maintained	Arbutus 'Marina' : Hybrid Strawberry Tree	2547 Vallejo St	4	Sidewalk: Curb side : Cutout	Tree	Private
3	DPW Maintained	Afroparus gracilior : Fern Pine	100 Kissling St	3	:	Tree	Private
4	DPW Maintained	Afroparus gracilior : Fern Pine	100 Kissling St	1	:	Tree	Private
5	DPW Maintained	Afroparus gracilior : Fern Pine	100 Kissling St	4	:	Tree	Private
6	DPW Maintained	Arbutus 'Marina' : Hybrid Strawberry Tree	2547 Vallejo St	2	Sidewalk: Curb side : Cutout	Tree	Private
7	Undocumented	Thuja occidentalis 'Emerald' : Emerald Arborvitae	1670 09th Ave	1	Sidewalk: Curb side : Cutout	Landscaping	Private
8	Planning Code 138.1 required	Magnolia grandiflora 'Little Gem' : Little Gem Ma...	99x Tandang Sora	1	Sidewalk: Curb side : Cutout	Tree	Private
9	Planning Code 138.1 required	Magnolia grandiflora 'Little Gem' : Little Gem Ma...	99x Tandang Sora	3	Sidewalk: Curb side : Cutout	Tree	Private

Alteryx Designer x64 - SF_Trees_DataProfiling.yxmd

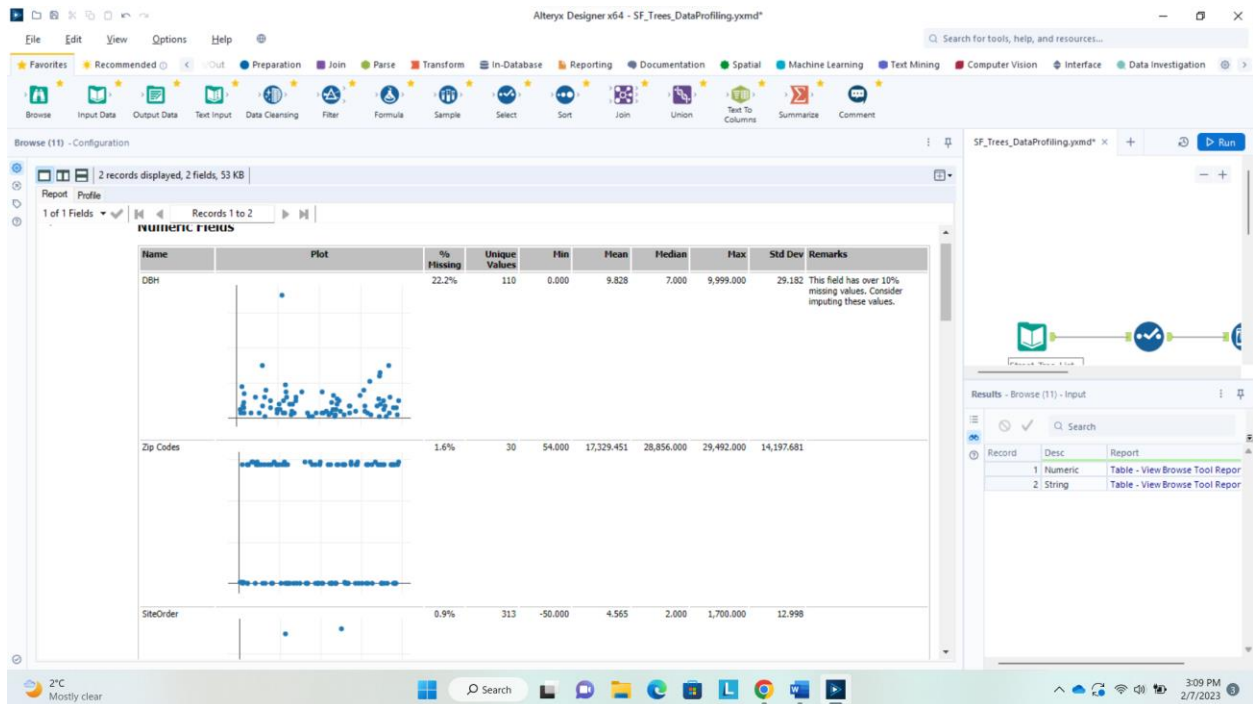
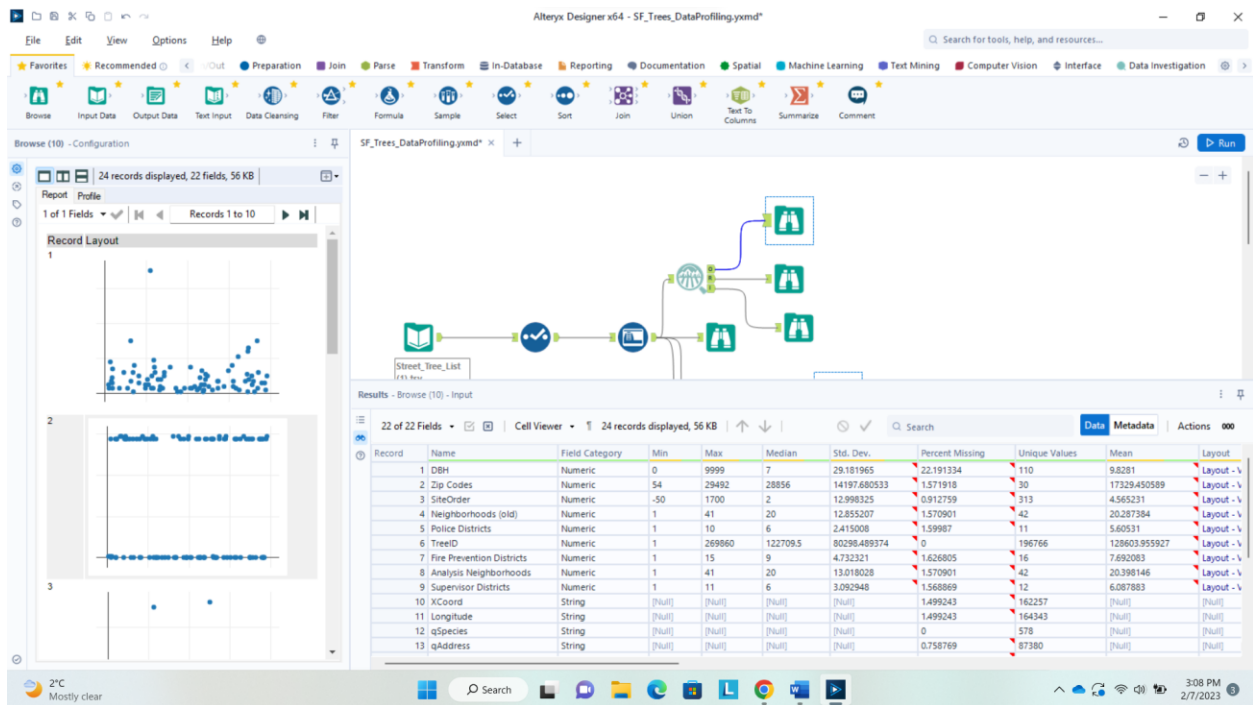
Browse (8) - Configuration

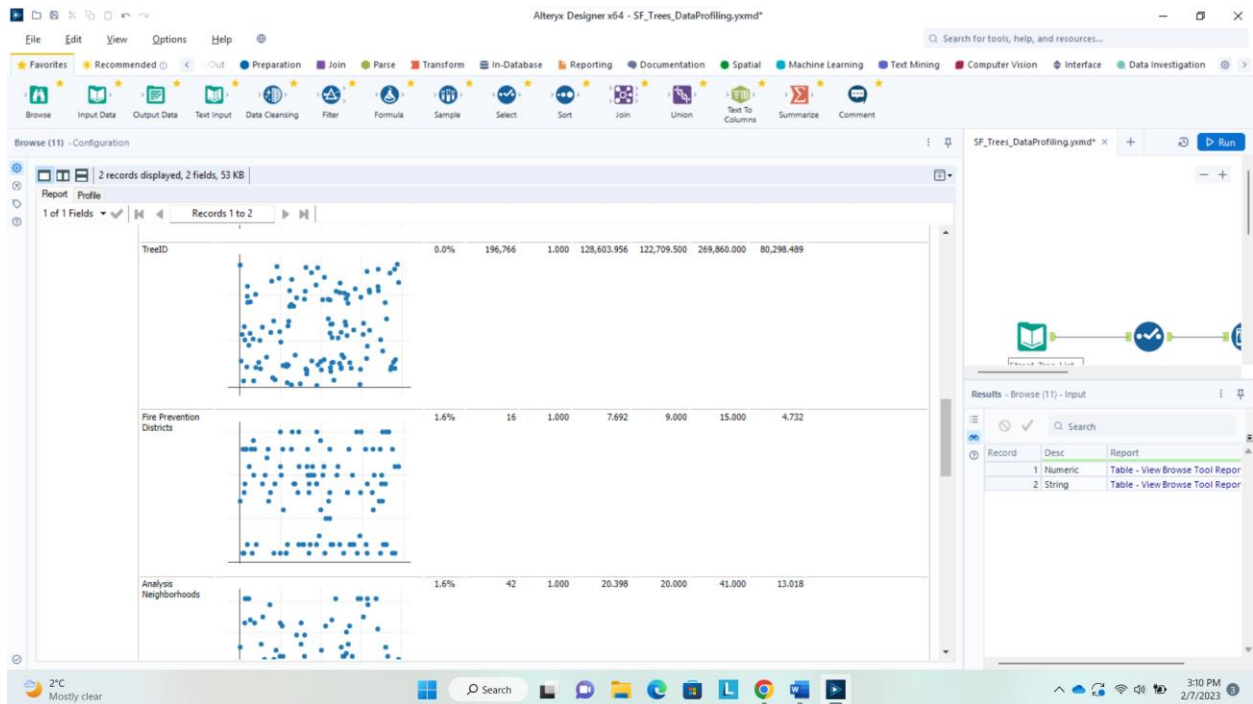
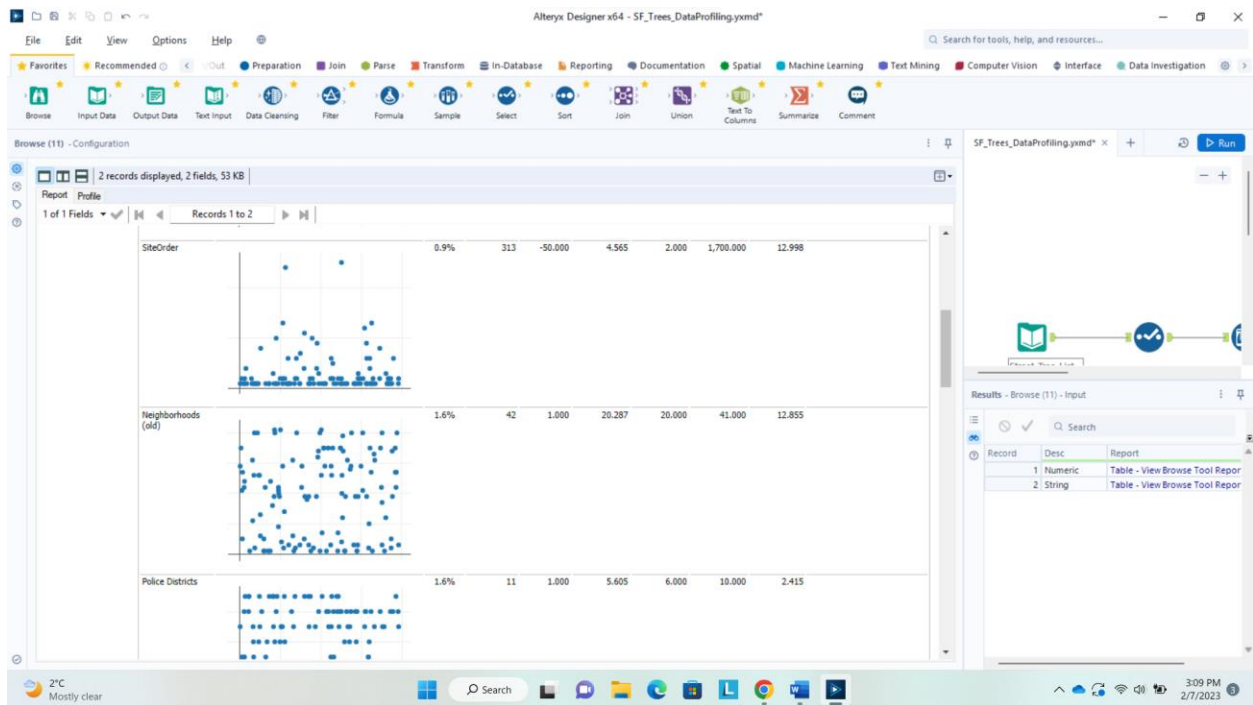
Profile: 196,766 records displayed, 25 fields, 23 MB

Workflow Diagram: Street_Tree_List (1).tsv -> Join -> Formula (7) -> Output Column: DI_Job_Name

Results - Browse (8) - Input

Record	Fire Prevention Districts	Police Districts	Supervisor Districts	Zip Codes	Neighborhoods (old)	Analysis Neighborhoods	DI_Job_Name
1	13	9	1	57	27	30	SF_Trees_DataProfiling.yxmd
2	13	9	1	57	27	30	SF_Trees_DataProfiling.yxmd
3	8	2	9	28853	19	20	SF_Trees_DataProfiling.yxmd
4	8	2	9	28853	19	20	SF_Trees_DataProfiling.yxmd
5	8	2	9	28853	19	20	SF_Trees_DataProfiling.yxmd
6	13	9	1	57	27	30	SF_Trees_DataProfiling.yxmd
7	1	8	4	56	12	14	SF_Trees_DataProfiling.yxmd
8	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	SF_Trees_DataProfiling.yxmd
9	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	SF_Trees_DataProfiling.yxmd
10	15	5	5	29492	9	3	SF_Trees_DataProfiling.yxmd
11	15	5	5	29492	9	3	SF_Trees_DataProfiling.yxmd
12	15	5	5	28862	3	5	SF_Trees_DataProfiling.yxmd
13	14	2	9	28856	34	34	SF_Trees_DataProfiling.yxmd
14	15	5	11	29492	9	3	SF_Trees_DataProfiling.yxmd
15	13	9	1	57	27	30	SF_Trees_DataProfiling.yxmd
16	13	9	1	29490	27	30	SF_Trees_DataProfiling.yxmd
17	13	9	1	57	27	30	SF_Trees_DataProfiling.yxmd
18	13	9	1	57	27	30	SF_Trees_DataProfiling.yxmd





Alteryx Designer x64 - SF_Trees_DataProfiling.yxmd

File Edit View Options Help

Search for tools, help, and resources...

Favorites Recommended Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Browse (11) - Configuration

Report Profile

1 of 1 Fields 2 records displayed, 2 fields, 53 KB

Name	% Missing	Unique Values	Shortest Value	Longest Value	Min Value Count	Max Value Count	Remarks
XCoord	1.5%	162,257	5993451	6012022.8946024915	1	2,950	Some values of this field have a small number of value counts. If appropriate, consider combining some value levels together.
Longitude	1.5%	164,343	-122.40714427	-122.4398698930563	1	2,950	Some values of this field have a small number of value counts. If appropriate, consider combining some value levels together.
qSpecies	0.0%	578	::	Leptospermum scoparium Yelene Strybing :: Helene Strybing New Zealand Tea Tree	1	11,811	Some values of this field have a small number of value counts. If appropriate, consider combining some value levels together.
qAddress	0.8%	87,380	0	525or larger Golden Gate Ave In A 24 Box	1	1,493	Some values of this field have a small number of value counts. If appropriate, consider combining some value levels together.
PermitNotes	72.8%	35,637	12	owner Victor Seeto signed agreement 2/10/10 to maintain bricks in tree basin in a safe/accessible manner that does not damage tree or restrict growth.	1	143,296	This field has over 10% missing values. Consider imputing these values. Some values of this field have a small number of value counts. If appropriate, consider combining some value levels together.
qCaretaker	0.0%	26	DPW	Landscape Architecture Capital Project	1	164,712	Some values of this field have a small number of value counts. If appropriate, consider combining some value levels together.
qCareAssistant	87.4%	20	FUF	Clary Bros. Landscape	1	172,052	This field has over 10% missing values. Consider imputing these values. Some values of this field have a small number of value counts. If appropriate, consider combining some value levels together.
PlotSize	25.6%	607	9	within landscaping plot	1	50,379	This field has over 10% missing values. Consider imputing these values. Some values of this field have a small number of value counts. If appropriate, consider combining some value levels together.
Latitude	1.5%	164,340	37.778228754	37.794558498932695	1	2,950	Some values of this field have a small number of value counts. If appropriate, consider combining some value levels together.
V_Pound	1.6%	143,791	3164133	396070.4813473444	1	3,600	Some values of this field have a small number of value counts. If appropriate, consider combining some value levels together.

2°C Mostly clear

Alteryx Designer x64 - SF_Trees_DataProfiling.yxmd

File Edit View Options Help

Search for tools, help, and resources...

Favorites Recommended Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Browse (12) - Configuration

Report Profile

1 of 1 Fields 1 record displayed, 1 field, 136 KB

Field Summary

Select variables to view

Record Text

1

Zip Codes

SiteOrder

Neighborhoods (old)

Police Districts

TreeID

Fire Prevention Districts

Analysis Neighborhoods

DBH

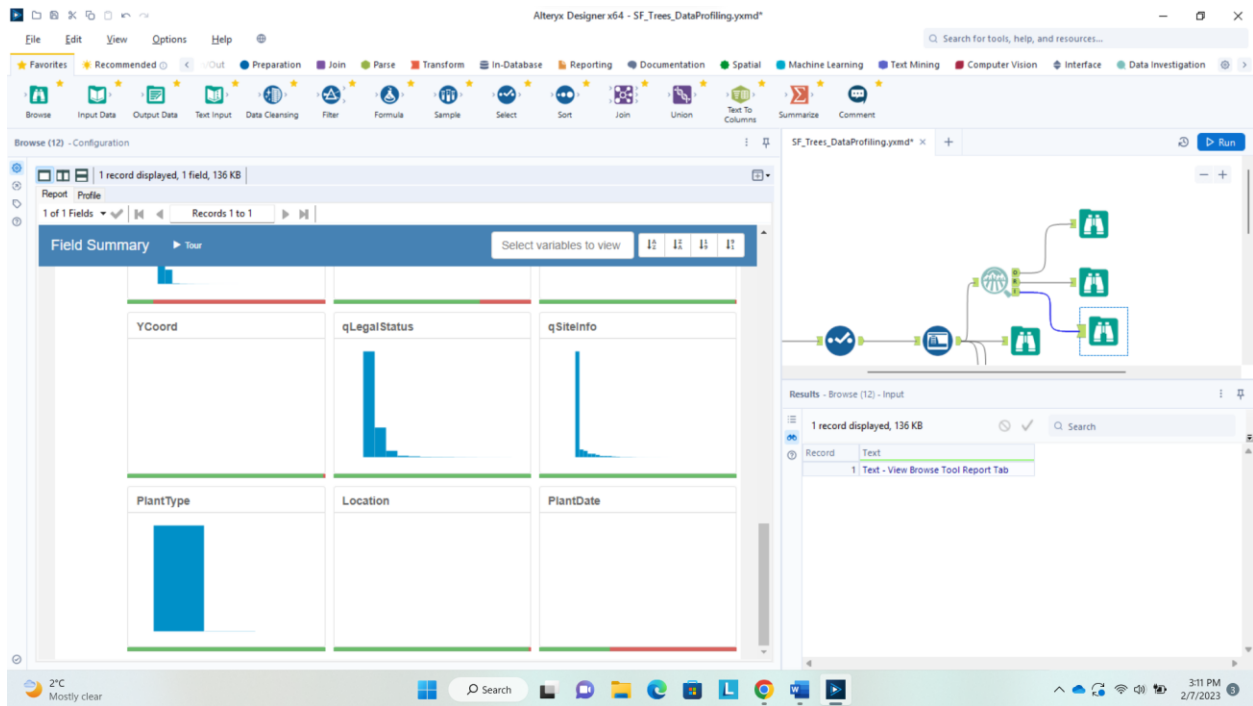
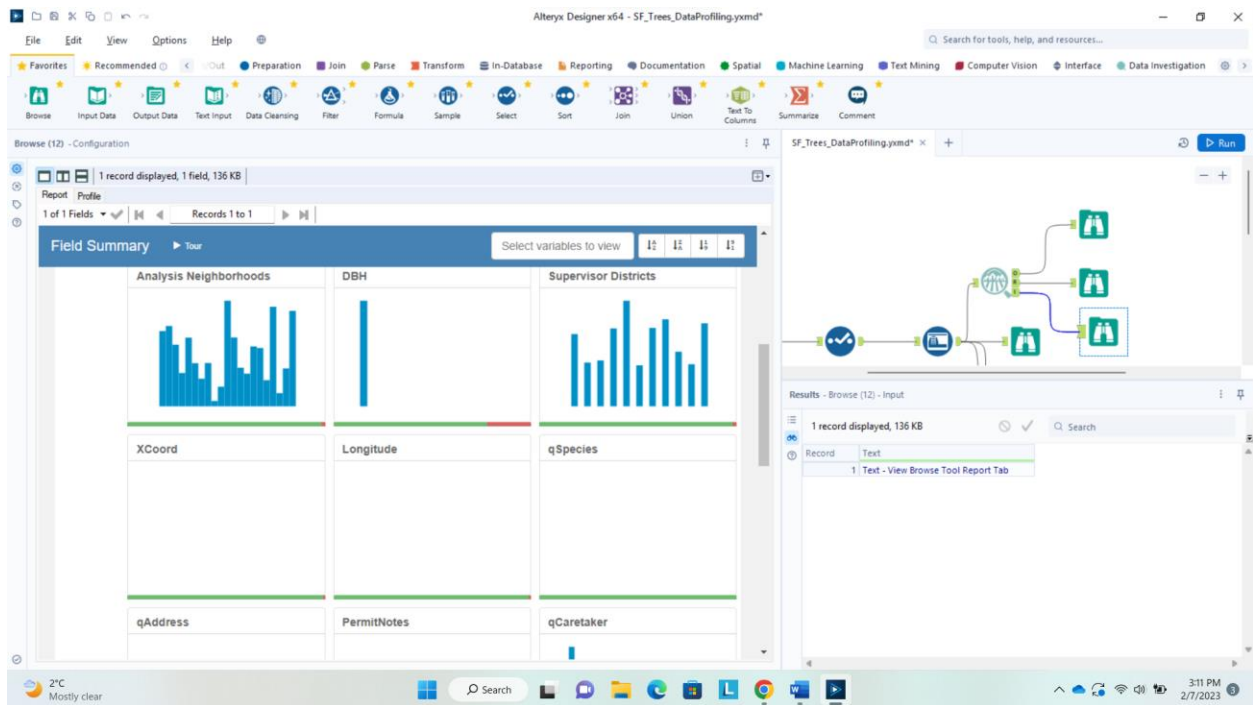
Supervisor Districts

Results - Browse (12) - Input

1 record displayed, 136 KB

Record	Text
1	Text - View Browse Tool Report Tab

2°C Mostly clear



Alteryx Designer x64 - SF_Trees_DataProfiling.yxmd*

File Edit View Options Help

Search for tools, help, and resources...

Favorites Recommended Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Browse Input Data Output Data Text Input Data Cleansing Filter Formula Sample Select Sort Join Union Text To Column Summarize Comment

Browse (14) - Configuration

Profile

531 records displayed, 3 fields, 12 KB

531 records displayed, 3 fields, 12 KB

Field Name

Analysis Neighborhoods 24

DBH 24

Fire Prevention Districts 24

Neighborhoods (old) 24

Police Districts 24

19 more >

Name

Data Type 24

Description 24

Maximum 24

Minimum 24

Name 24

28 more >

Value

0 68

File: C:\Users\moins\Downloads\Street_Tree_... 24

1 14

SF_Trees_DataProfiling.yxmd*

Run

Results - Browse (14) - Input

3 of 3 Fields Cell Viewer 531 records displayed, 12 KB

Record	FieldName	Name	Value
1	TreeID	Name	TreeID
2	TreeID	Data Type	Int32
3	TreeID	Size	4
4	TreeID	Source	File: C:\Users\moins\Downloads\Street_Tree_List (...)
5	TreeID	Description	{Null}
6	TreeID	OKs	196766
7	TreeID	Nulls	0
8	TreeID	Non-Nulls	196766
9	TreeID	Minimum	1
10	TreeID	Maximum	269860
11	TreeID	Average	128603.955927345
12	TreeID	Sum	2530485992
13	TreeID	Standard Deviation	80298.4893742973
14	TreeID	Variance	6447947395.79414
15	TreeID	Uniques	> 10000
16	TreeID	Unique Values	{Null}
17	TreeID	25th Percentile	53543.277
18	TreeID	25th Percentile Margin Of Error	-23.350, +3.632

2°C Mostly clear

Search

3:14 PM 2/7/2023