

# DC v. Next

Mid-Fi: Metrics Tab

## DESIGN TEAM

Alex Swain, Dillon Eversman, Stevie Johnston, Nishita Muhnot,  
Voranoth Supadulya

05/10/17



# Metrics: All Columns Mode Interactions

## ROW

On-hover over a row-

Move, Hide: For all rows

Expand: Only for frequency row

## COLUMN

On-hover over a column-

Maximize/Minimize

## FREQUENCY SECTION

View All: when no. of bins > 20

On-hover over a graph-

Change bin count: For Number and Date data types only

Inverse/Reverse: For all data types

## PATTERNS SECTION

More: when no. of patterns > 1

Select an operation or enter a command to cleanse and shape your data



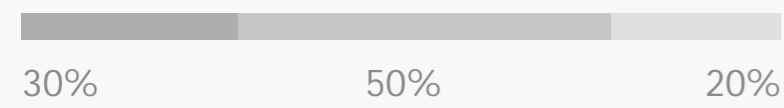
Data Visualizations Metrics Suggestions Data Flow

COLUMN NAME  
Number

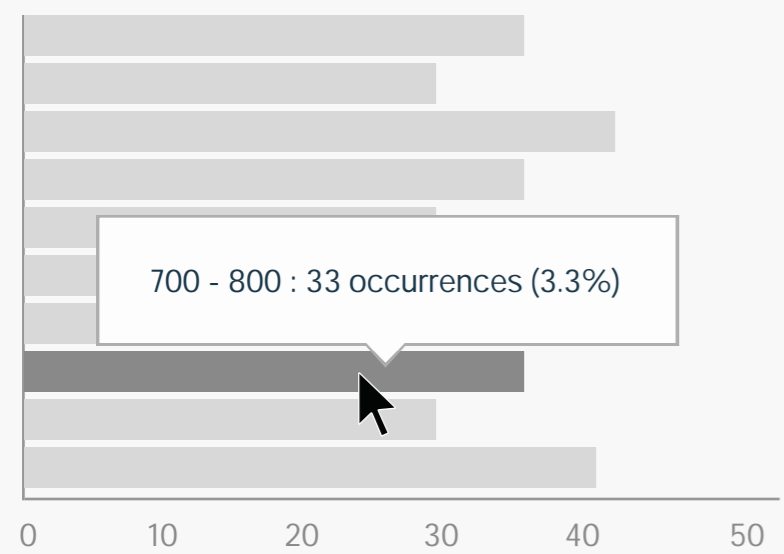


COLUMN NAME  
String

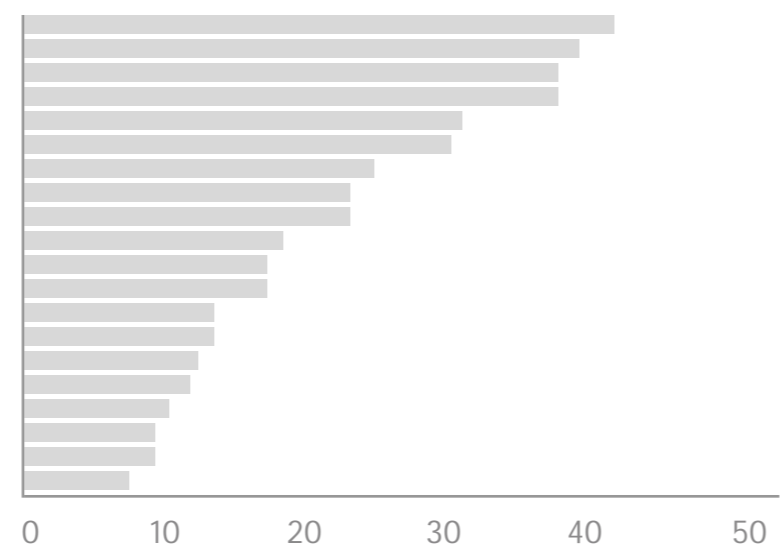
COLUMN NAME  
Date



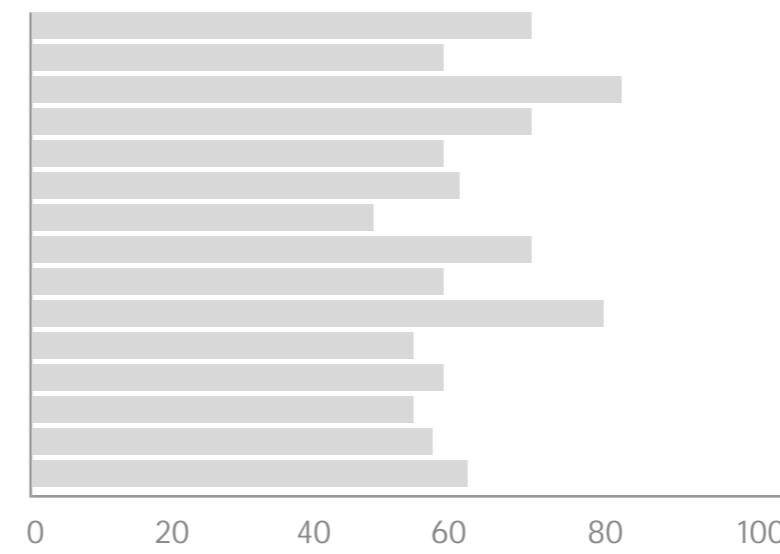
FREQUENCY Bins: 10



FREQUENCY



FREQUENCY



View All

STATISTICS

Minimum	1.00
Lower Quartile	250.50
Median	500.00
Upper Quartile	750.50
Maximum	1000.00
Average	500.33
Outliers	5

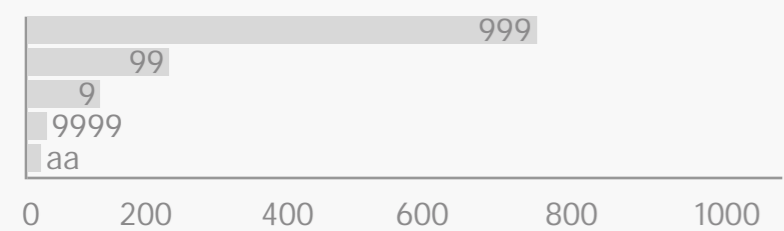
STATISTICS

Unique	50
String Length Outliers	1
Minimum Length	2
Maximum Length	11
Average Length	7

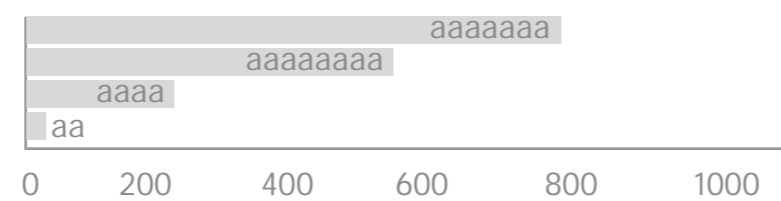
STATISTICS

Minimum	Jan 17 2001 00:00
Median	Aug 18 2008 00:00
Average	Jun 17 2008 06:25
Maximum	Dec 8 2015 00:00
Outliers	1

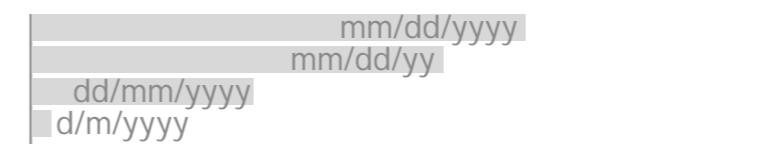
PATTERNS



PATTERNS



PATTERNS



More



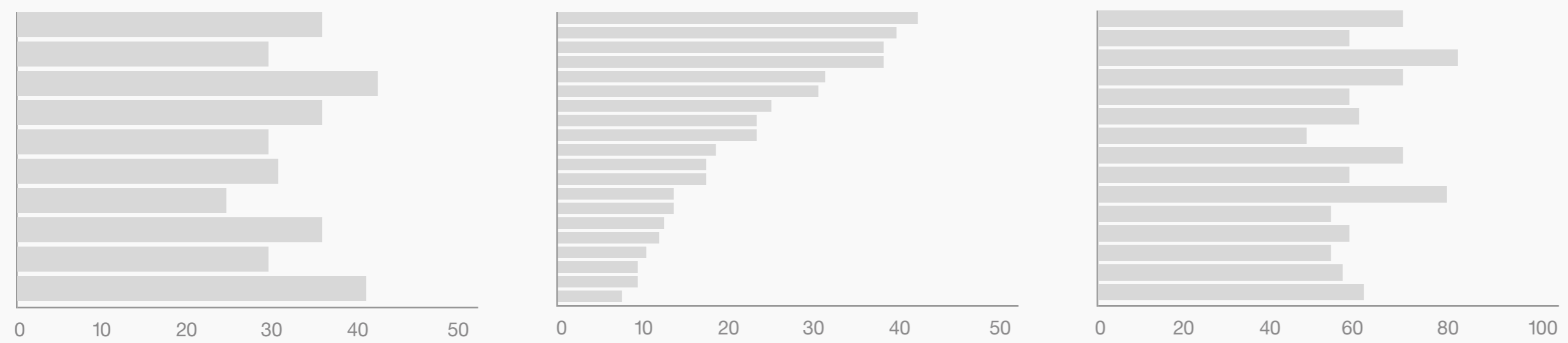
⊕ Select an operation or enter a command to cleanse and shape your data 🔍

Data Visualizations Metrics Suggestions Data Flow

<b>COLUMN NAME</b> Number	<b>COLUMN NAME</b> String	<b>COLUMN NAME</b> Date
------------------------------	------------------------------	----------------------------



FREQUENCY FREQUENCY FREQUENCY

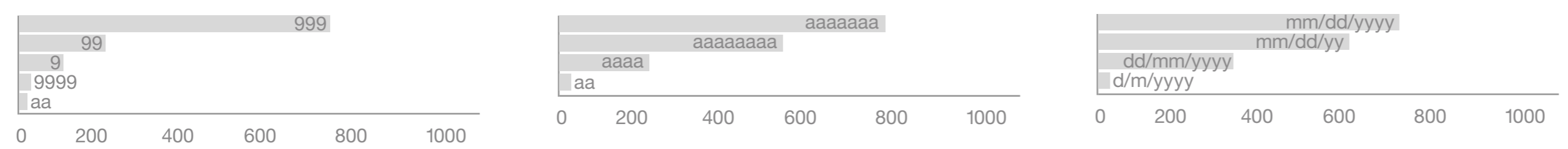


View All

STATISTICS STATISTICS STATISTICS

Minimum	1.00	Unique	50	Minimum	Jan 17 2001 00.00
Lower Quartile	250.50	Minimum Length	2	Median	Aug 18 2008 00.00
Median	500.00	Maximum Length	11	Average	Jun 17 2008 06:25
Average	500.33	Average Length	7	Maximum	Dec 8 2015 00.00
Upper Quartile	750.50	String Length Outliers	1	Outliers	1
Maximum	1000.00				
Outliers	5				

PATTERNS PATTERNS PATTERNS



More





Select an operation or enter a command to cleanse and shape your data



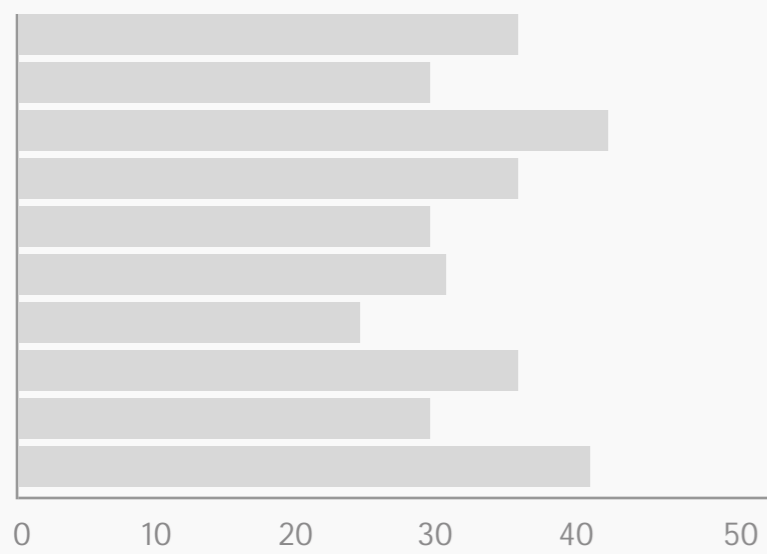
Data Visualizations Metrics Suggestions Data Flow

COLUMN NAME  
Number

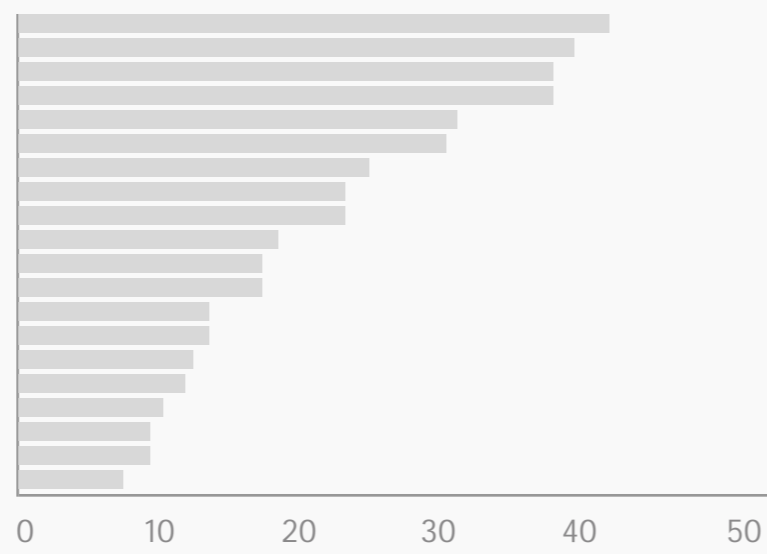
COLUMN NAME  
String

COLUMN NAME  
Date

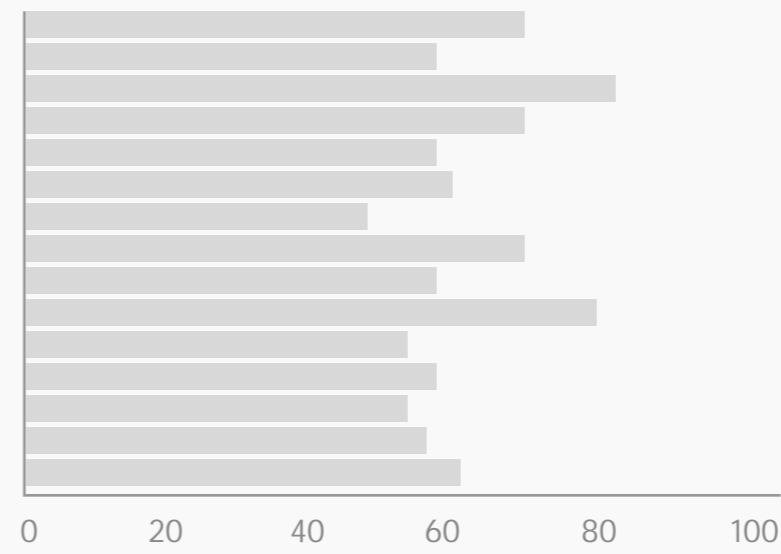
FREQUENCY



FREQUENCY



FREQUENCY



View All

STATISTICS

Minimum	1.00
Lower Quartile	250.50
Median	500.00
Average	500.33
Upper Quartile	750.50
Maximum	1000.00
Outliers	5

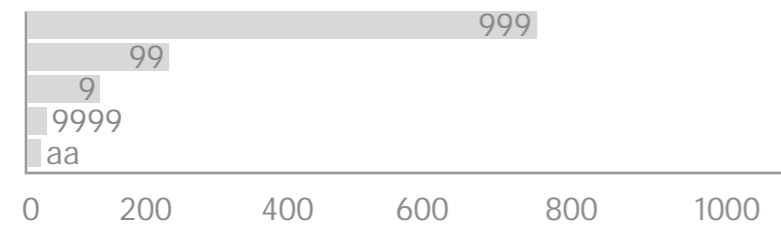
STATISTICS

Unique	50
Minimum Length	2
Maximum Length	11
Average Length	7
String Length Outliers	1

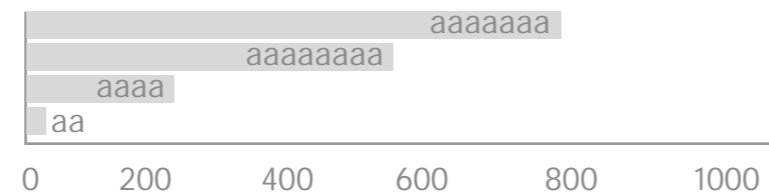
STATISTICS

Minimum	Jan 17 2001 00:00
Median	Aug 18 2008 00:00
Average	Jun 17 2008 06:25
Maximum	Dec 8 2015 00:00
Outliers	1

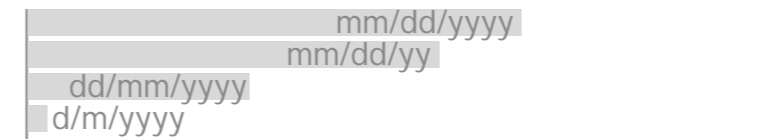
PATTERNS



PATTERNS



PATTERNS



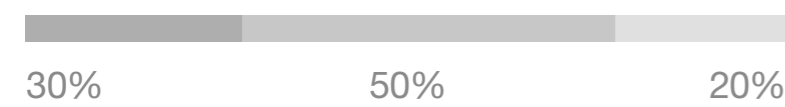
More



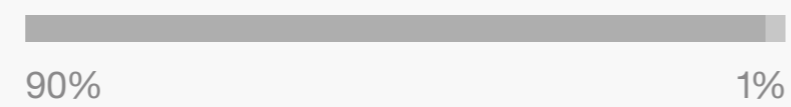
⊕ Select an operation or enter a command to cleanse and shape your data 🔍

Data Visualizations Metrics Suggestions Data Flow

COLUMN NAME  
Number



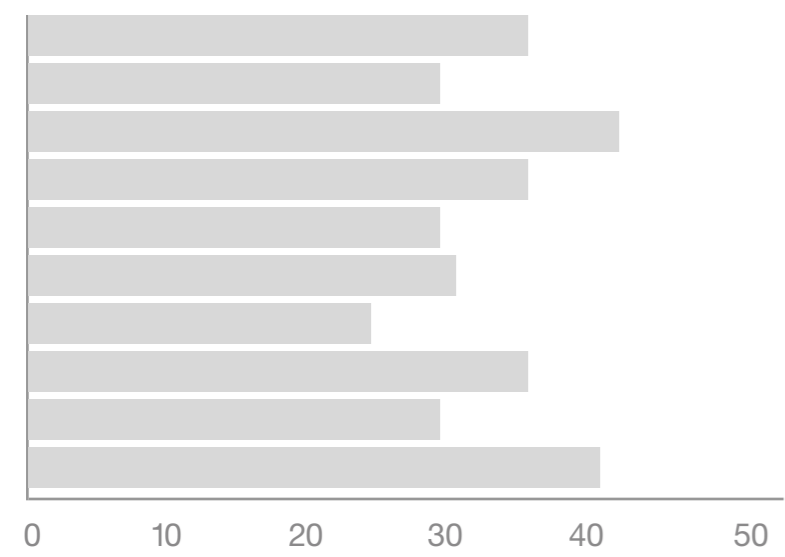
COLUMN NAME  
String



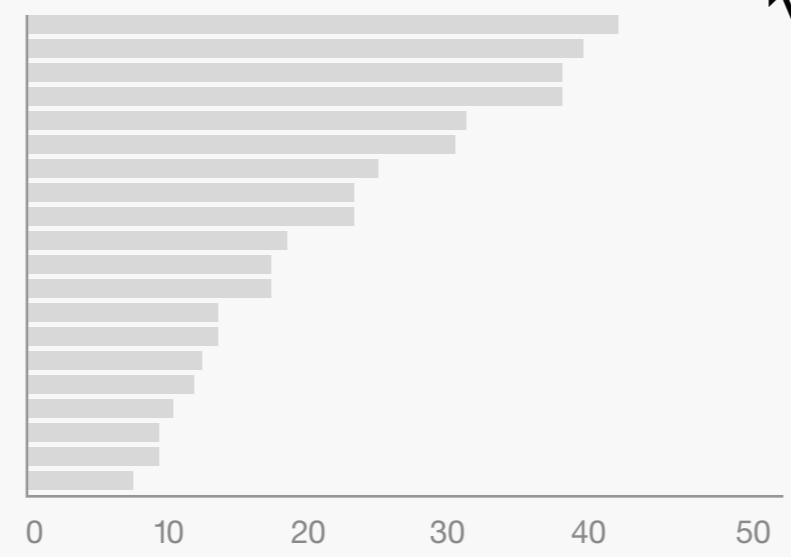
COLUMN NAME  
Date



FREQUENCY

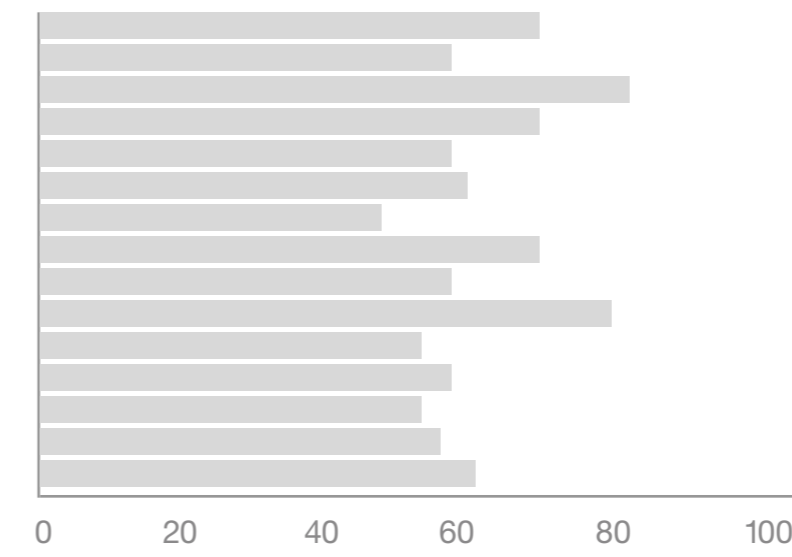


FREQUENCY



View All

FREQUENCY



STATISTICS

Minimum	1.00
Lower Quartile	250.50
Median	500.00
Average	500.33
Upper Quartile	750.50
Maximum	1000.00
Outliers	5

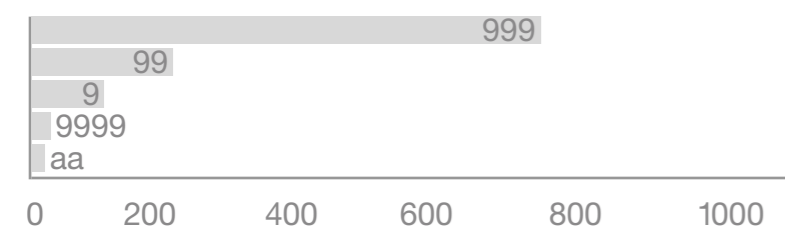
STATISTICS

Unique	50
Minimum Length	2
Maximum Length	11
Average Length	7
String Length Outliers	1

STATISTICS

Minimum	Jan 17 2001 00:00
Median	Aug 18 2008 00:00
Average	Jun 17 2008 06:25
Maximum	Dec 8 2015 00:00
Outliers	1

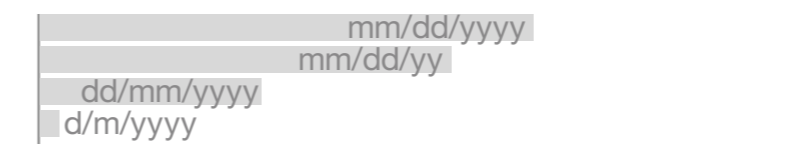
PATTERNS



PATTERNS



PATTERNS



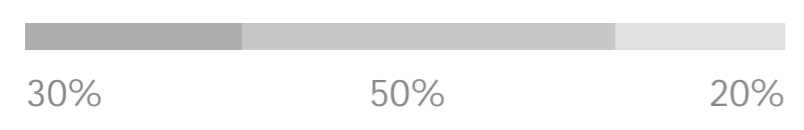
More



⊕ Select an operation or enter a command to cleanse and shape your data 🔍

Data Visualizations Metrics Suggestions Data Flow

⋮  
COLUMN NAME  
Number



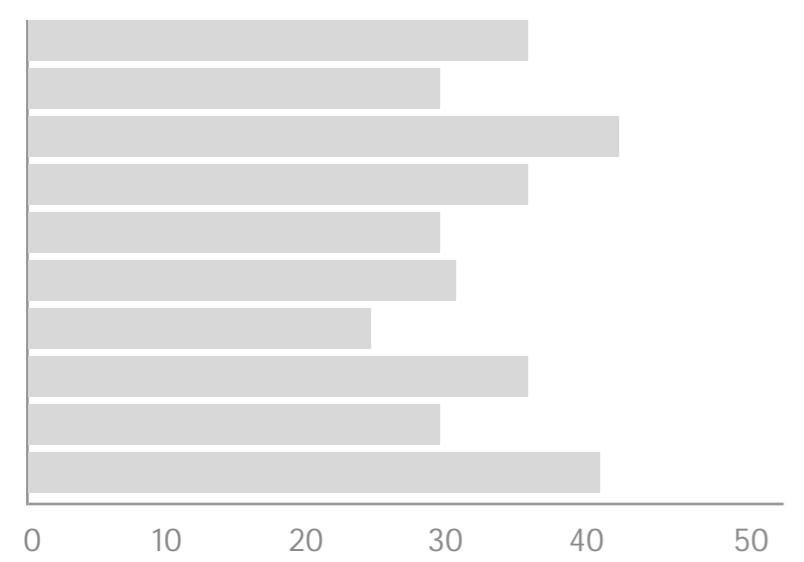
↗  
COLUMN NAME  
String



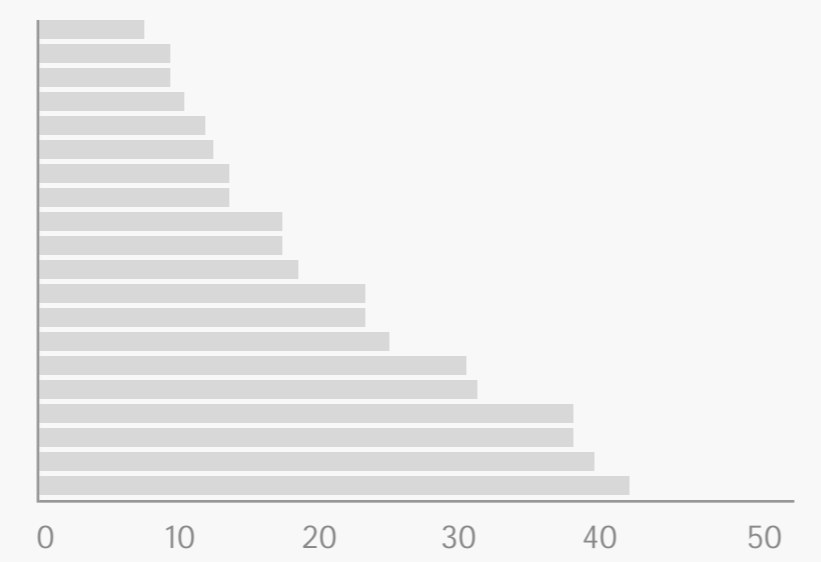
COLUMN NAME  
Date



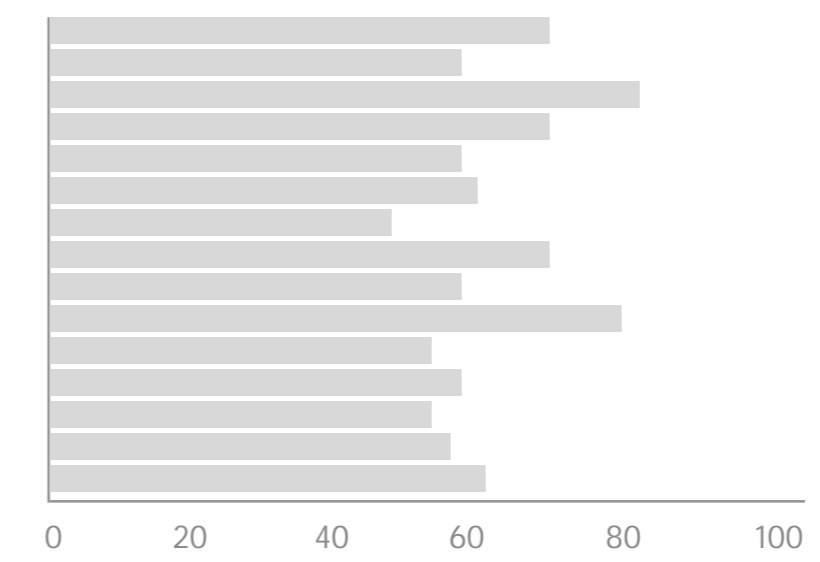
⏏  
FREQUENCY



↕  
FREQUENCY



FREQUENCY



STATISTICS

Minimum	1.00
Lower Quartile	250.50
Median	500.00
Average	500.33
Upper Quartile	750.50
Maximum	1000.00
Outliers	5

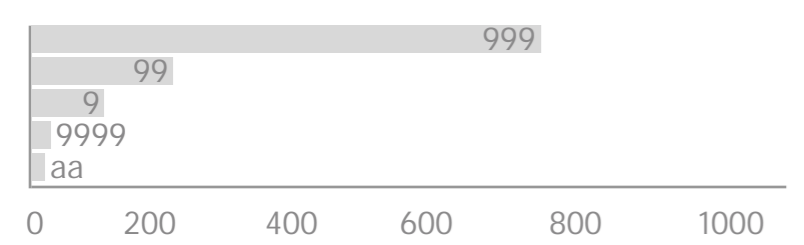
STATISTICS

Unique	50
Minimum Length	2
Maximum Length	11
Average Length	7
String Length Outliers	1

STATISTICS

Minimum	Jan 17 2001 00:00
Median	Aug 18 2008 00:00
Average	Jun 17 2008 06:25
Maximum	Dec 8 2015 00:00
Outliers	1

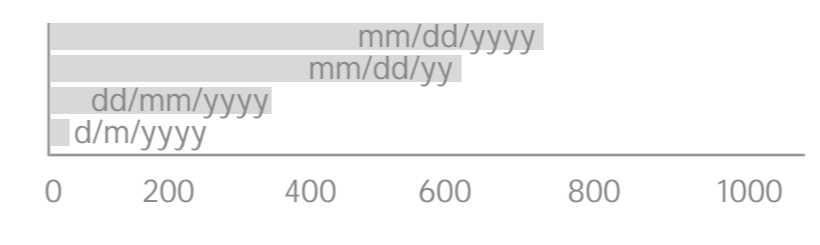
PATTERNS



PATTERNS



PATTERNS



More



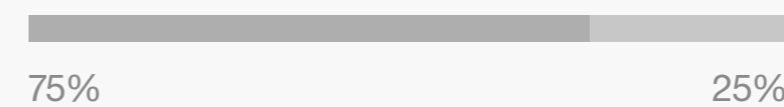
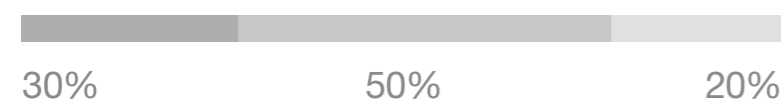
Select an operation or enter a command to cleanse and shape your data

Data Visualizations Metrics Suggestions Data Flow

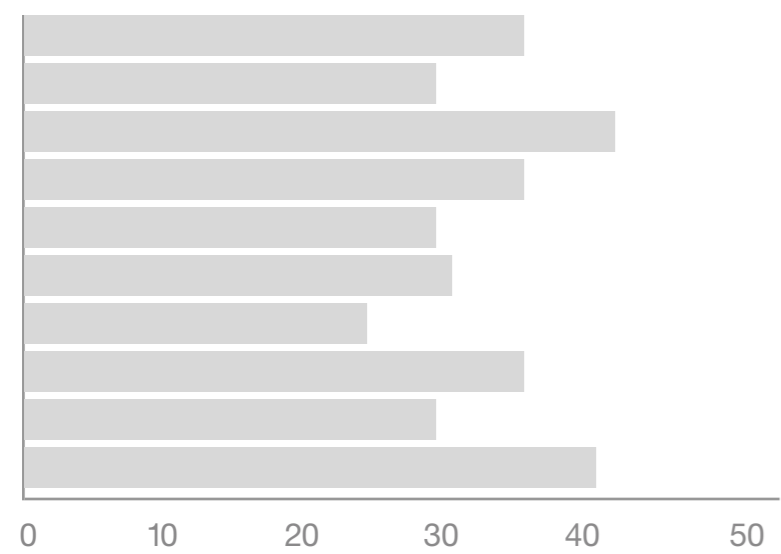
COLUMN NAME  
Number

COLUMN NAME  
String

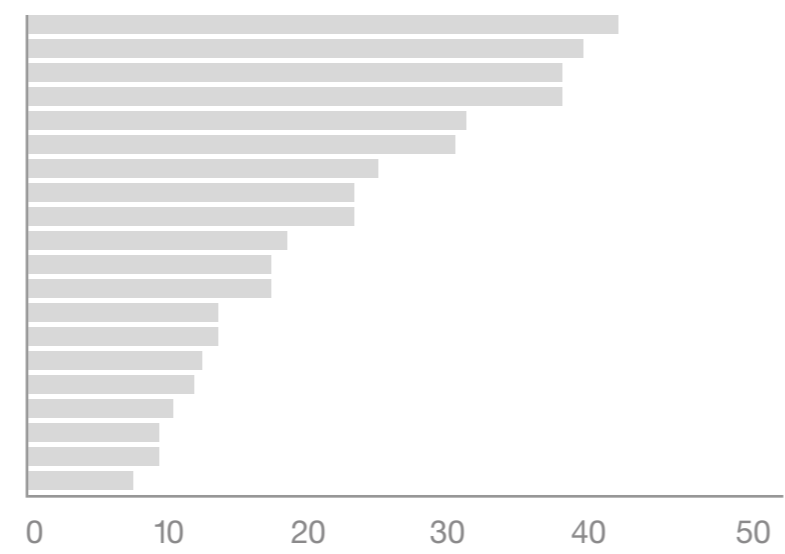
COLUMN NAME  
Date



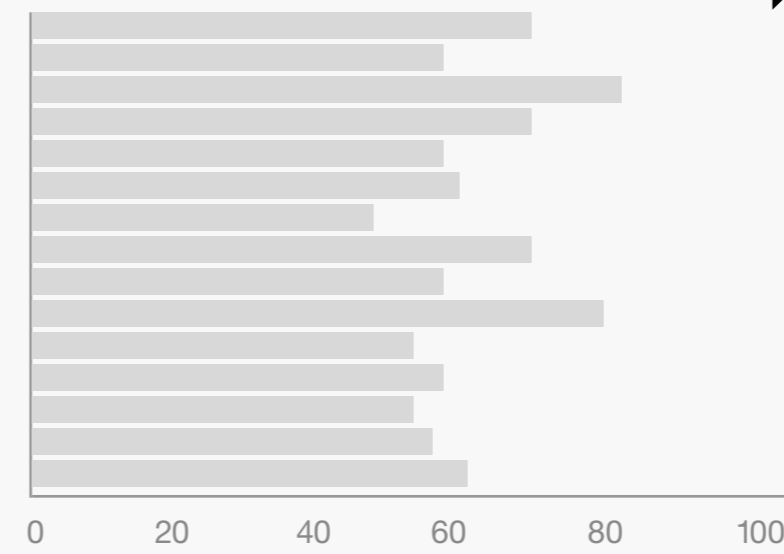
FREQUENCY



FREQUENCY



FREQUENCY



Bins :

View All

STATISTICS

Minimum	1.00
Lower Quartile	250.50
Median	500.00
Average	500.33
Upper Quartile	750.50
Maximum	1000.00
Outliers	5

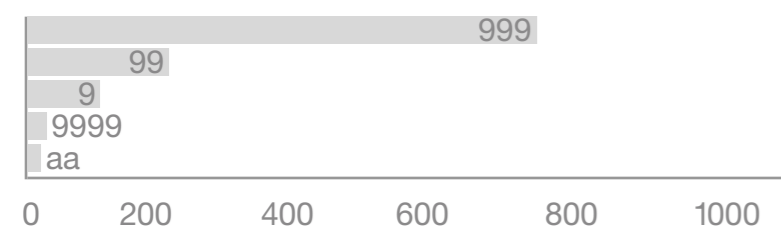
STATISTICS

Unique	50
Minimum Length	2
Maximum Length	11
Average Length	7
String Length Outliers	1

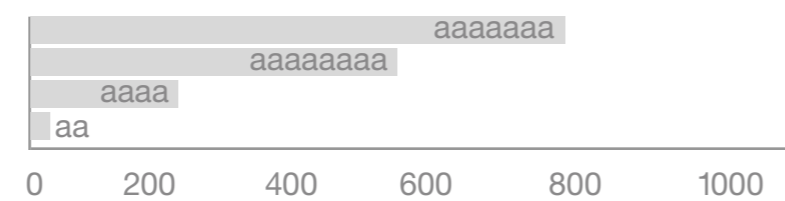
STATISTICS

Minimum	Jan 17 2001 00:00
Median	Aug 18 2008 00:00
Average	Jun 17 2008 06:25
Maximum	Dec 8 2015 00:00
Outliers	1

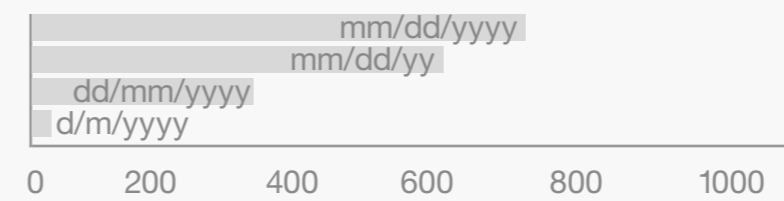
PATTERNS



PATTERNS



PATTERNS



More



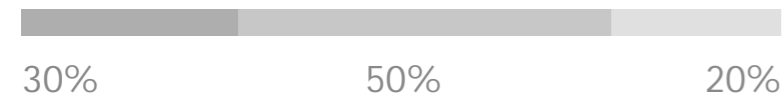
⊕ Select an operation or enter a command to cleanse and shape your data 🔍

Data Visualizations Metrics Suggestions Data Flow

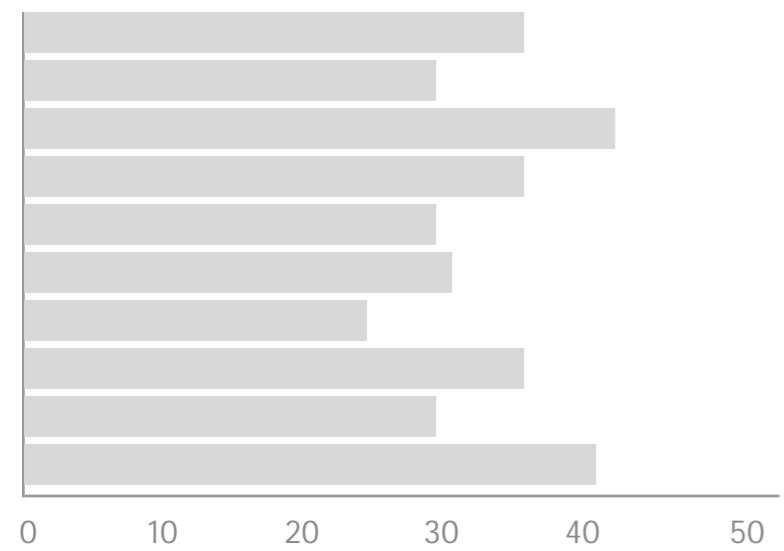
COLUMN NAME  
Number

COLUMN NAME  
String

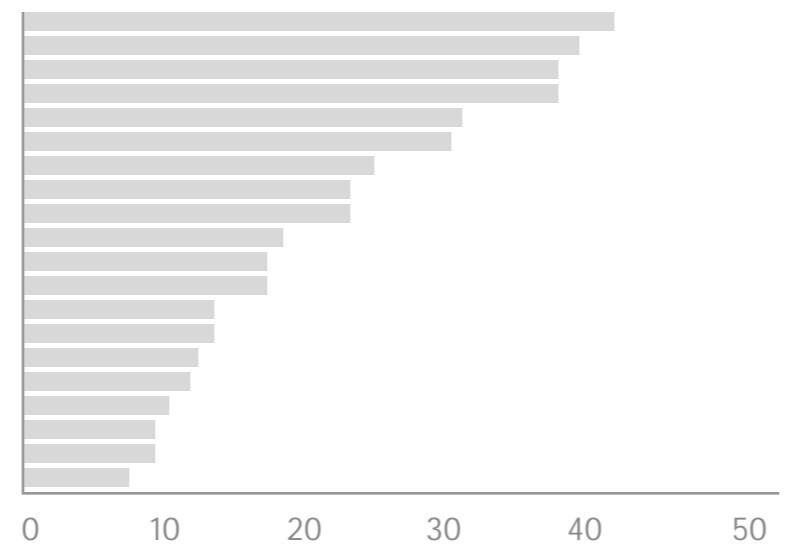
COLUMN NAME  
Date



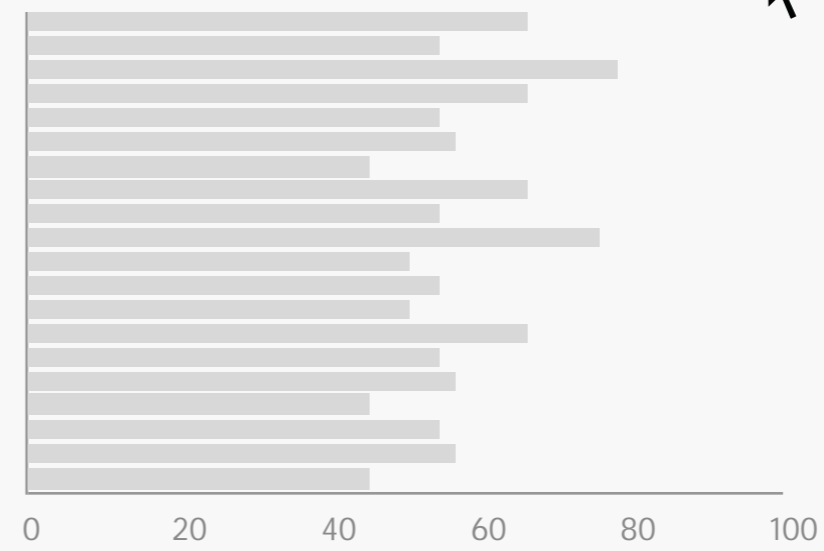
FREQUENCY



FREQUENCY



FREQUENCY



View All

View All

STATISTICS

Minimum	1.00
Lower Quartile	250.50
Median	500.00
Average	500.33
Upper Quartile	750.50
Maximum	1000.00
Outliers	5

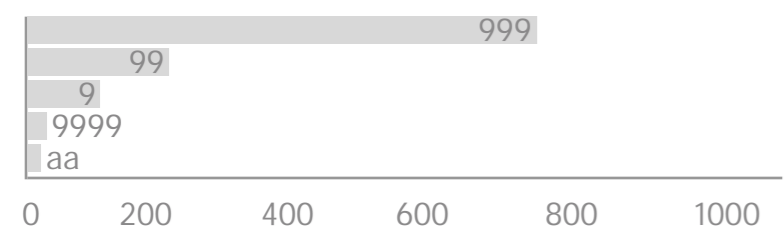
STATISTICS

Unique	50
Minimum Length	2
Maximum Length	11
Average Length	7
String Length Outliers	1

STATISTICS

Minimum	Jan 17 2001 00:00
Median	Aug 18 2008 00:00
Average	Jun 17 2008 06:25
Maximum	Dec 8 2015 00:00
Outliers	1

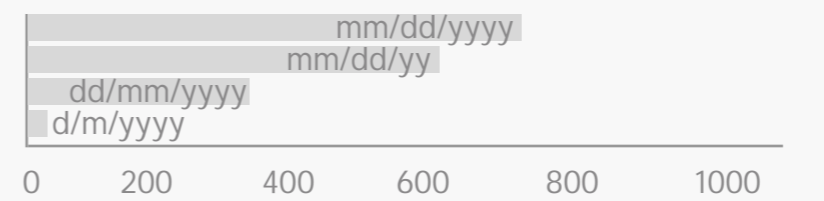
PATTERNS



PATTERNS



PATTERNS



More





Select an operation or enter a command to cleanse and shape your data

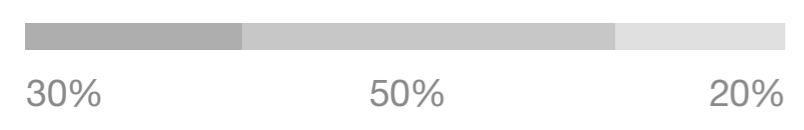


Data Visualizations Metrics Suggestions Data Flow

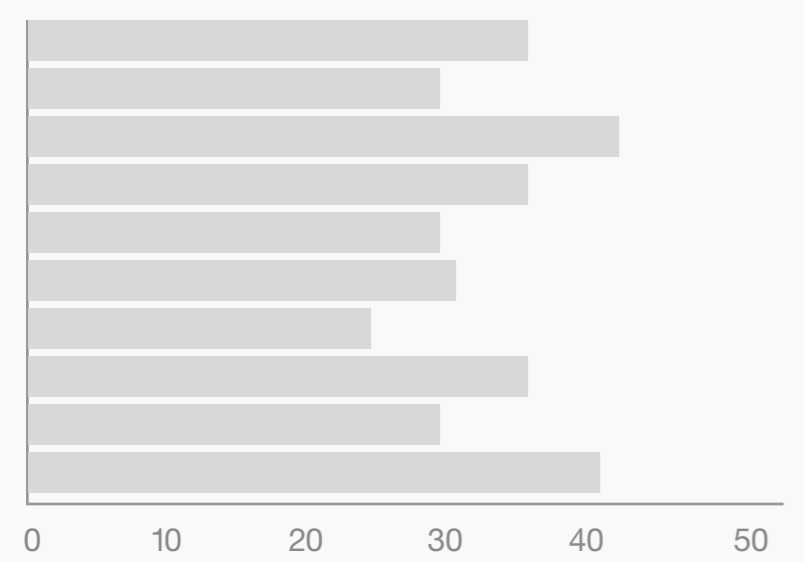
COLUMN NAME  
Number

COLUMN NAME  
String

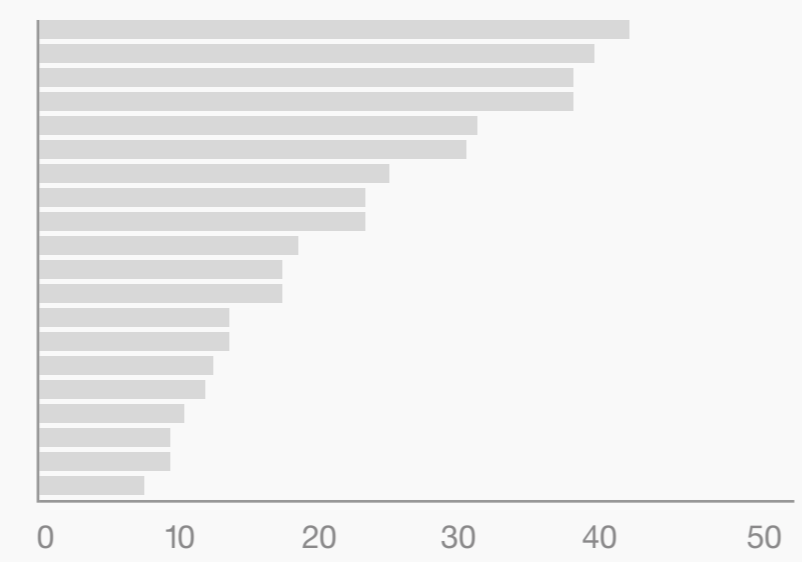
COLUMN NAME  
Date



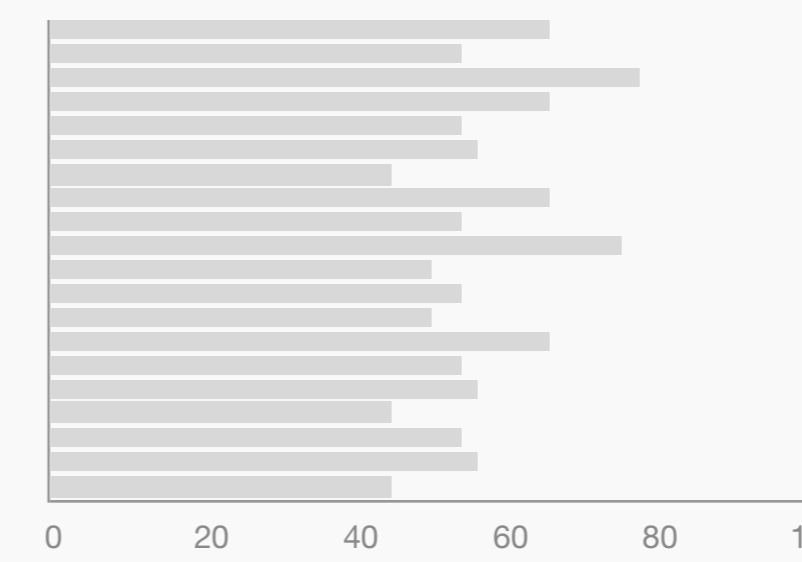
FREQUENCY



FREQUENCY



FREQUENCY



View All

View All

STATISTICS

Minimum	1.00
Lower Quartile	250.50
Median	500.00
Average	500.33
Upper Quartile	750.50
Maximum	1000.00
Outliers	5

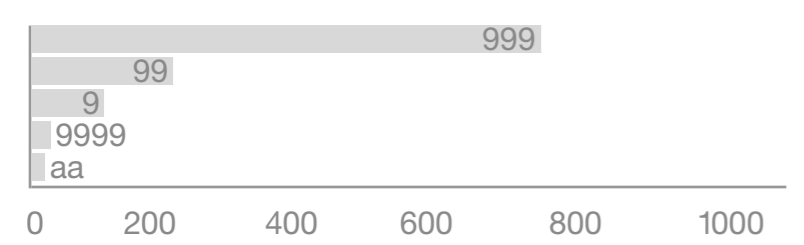
STATISTICS

Unique	50
Minimum Length	2
Maximum Length	11
Average Length	7
String Length Outliers	1

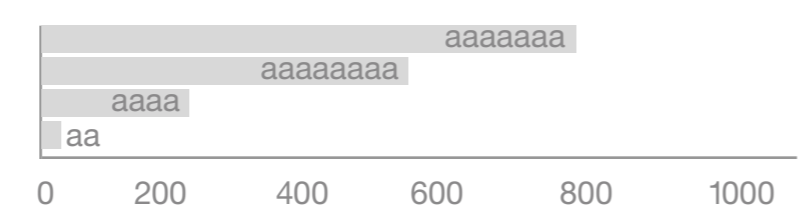
STATISTICS

Minimum	Jan 17 2001 00:00
Median	Aug 18 2008 00:00
Average	Jun 17 2008 06:25
Maximum	Dec 8 2015 00:00
Outliers	1

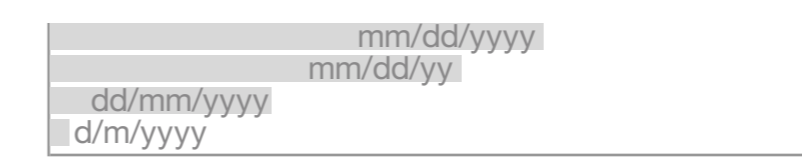
PATTERNS



PATTERNS



PATTERNS



More





Select an operation or enter a command to cleanse and shape your data



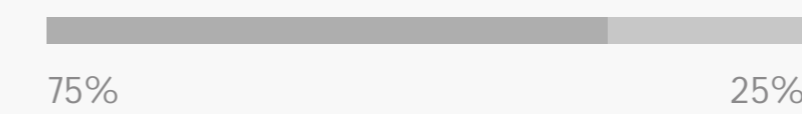
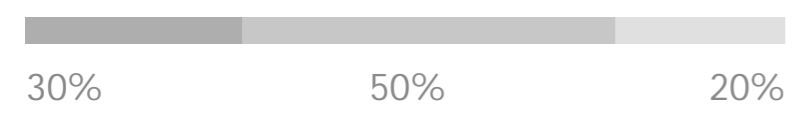
Data Visualizations Metrics

Suggestions Data Flow

COLUMN NAME  
Number

COLUMN NAME  
String

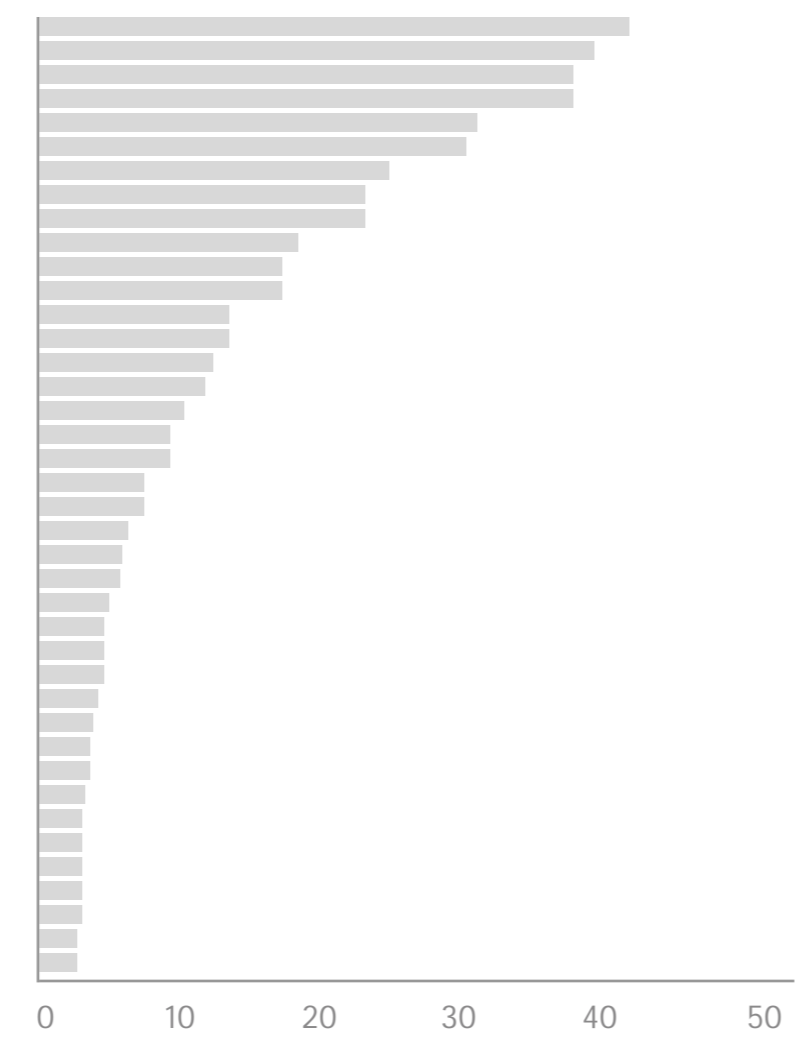
COLUMN NAME  
Date



FREQUENCY



FREQUENCY



FREQUENCY

Bins :



View All

View All

STATISTICS

Minimum	1.00
Lower Quartile	250.50
Median	500.00
Average	500.33
Upper Quartile	750.50
Maximum	1000.00
Outliers	5

STATISTICS

Unique	50
Minimum Length	2
Maximum Length	11
Average Length	7
String Length Outliers	1

STATISTICS

Minimum	Jan 17 2001 00:00
Median	Aug 18 2008 00:00
Average	Jun 17 2008 06:25
Maximum	Dec 8 2015 00:00
Outliers	1



⊕ Select an operation or enter a command to cleanse and shape your data 🔍

Data Visualizations Metrics Suggestions Data Flow

⋮ COLUMN NAME Date ↗️

### STATISTICS

Minimum	Jan 17 2001 00.00
Median	Aug 18 2008 00.00
Average	Jun 17 2008 06:25
Maximum	Dec 8 2015 00.00
Outliers	1

### DATA QUALITY

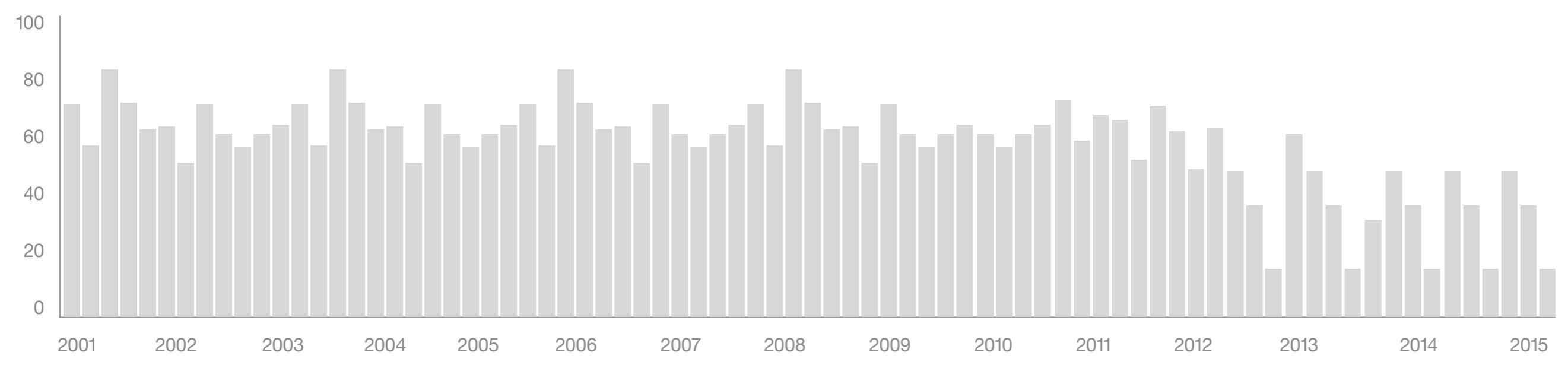
Matched	Mismatched
75%	25%
Unique	
100%	

### MISMATCHED VALUES AND OUTLIERS (258)

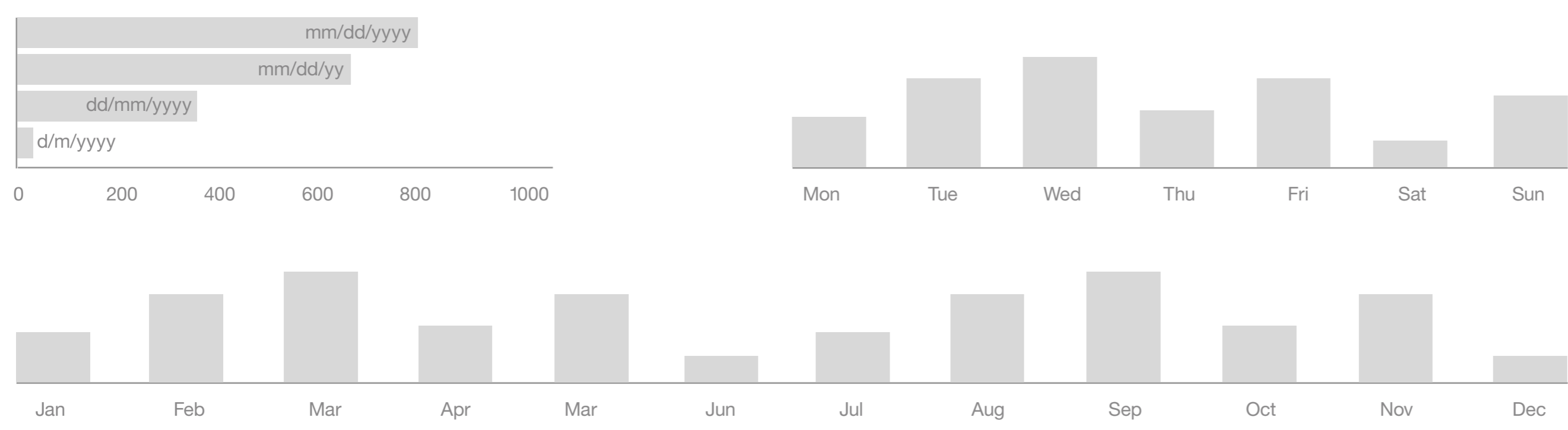
- 22/02/2015
- 1/9/2008
- 06/12/03
- 02-05-88
- 05-02-86

[View All](#)

### FREQUENCY



### PATTERNS





Select an operation or enter a command to cleanse and shape your data

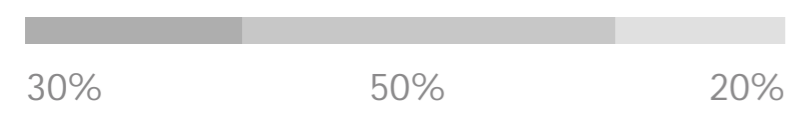


Data Visualizations Metrics Suggestions Data Flow

COLUMN NAME  
Number

COLUMN NAME  
String

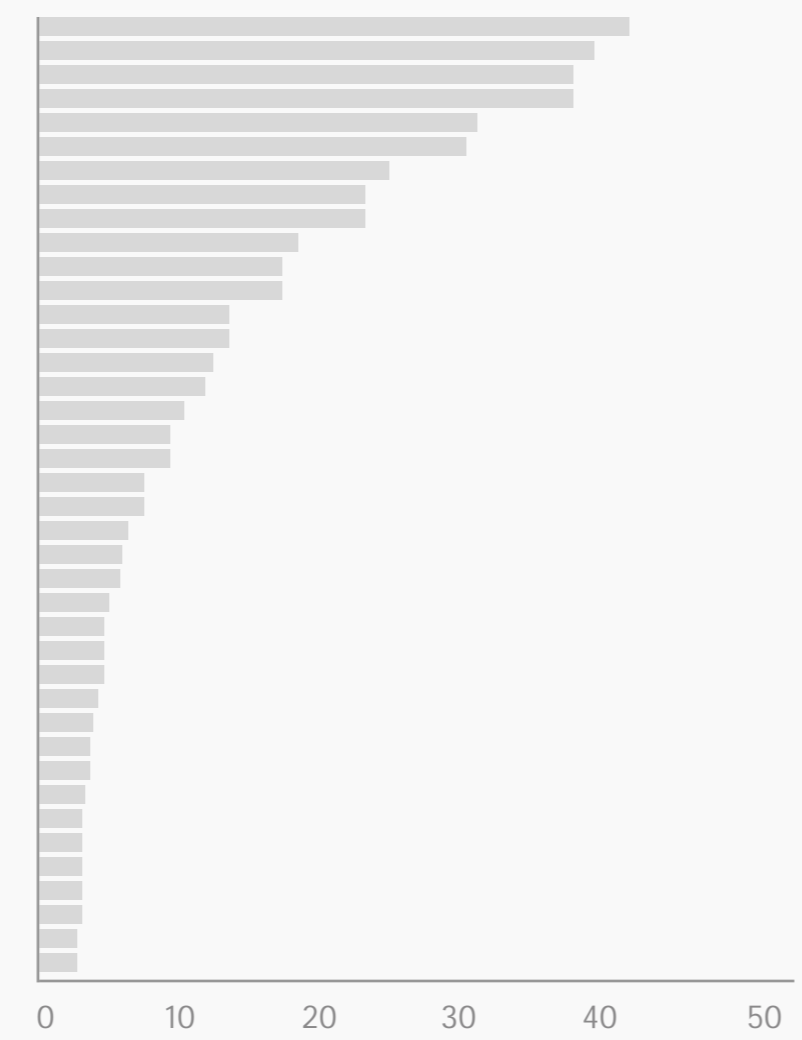
COLUMN NAME  
Date



FREQUENCY



FREQUENCY



FREQUENCY



View All

View All

STATISTICS

Minimum	1.00
Lower Quartile	250.50
Median	500.00
Average	500.33
Upper Quartile	750.50
Maximum	1000.00
Outliers	5

STATISTICS

Unique	50
Minimum Length	2
Maximum Length	11
Average Length	7
String Length Outliers	1

STATISTICS

Minimum	Jan 17 2001 00:00
Median	Aug 18 2008 00:00
Average	Jun 17 2008 06:25
Maximum	Dec 8 2015 00:00
Outliers	1





+ Select an operation or enter a command to cleanse and shape your data



Data Visualizations Metrics

Suggestions Data Flow



SELECTION MODE

Select table

Select all columns

Select certain columns

TYPE AND CLASSIFICATION

Data type

1.00

STATISTICS

Unique

50

✓ Classification

250.50

String Length Outliers

1

STATISTICS

Outliers

1

METRIC VISUALIZATIONS

500.00

Minimum Length

2

Minimum

Jan 17 2001 00.00

✓ Data quality

750.50

Maximum Length

11

Median

Aug 18 2008 00.00

Frequency

1000.00

Average Length

7

Maximum

Dec 8 2015 00.00

✓ Statistics

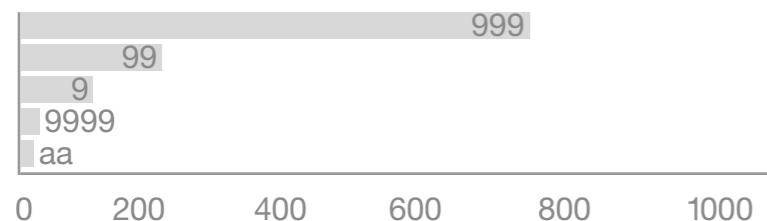
5

Average

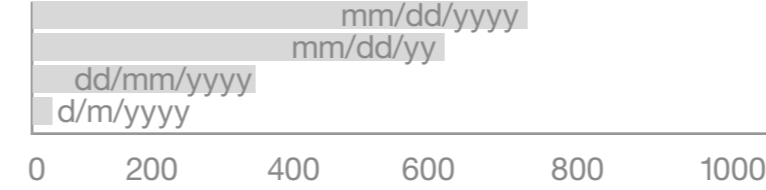
Jun 17 2008 06:25

✓ Patterns

PATTERNS



PATTERNS



More



# Metrics: Drill-in Mode Interactions

## **FREQUENCY SECTION**

On-hover over a graph-

Change bin count: For Number and Date data types only

Inverse/Reverse: For all data types

## **MISMATCHED VALUES AND OUTLIERS SECTION**

View All: when no. of values > 5

⊕ Select an operation or enter a command to cleanse and shape your data 🔍

Data Visualizations Metrics Suggestions Data Flow

⋮ COLUMN NAME Date ↗️

STATISTICS

Minimum	Jan 17 2001 00.00
Median	Aug 18 2008 00.00
Average	Jun 17 2008 06:25
Maximum	Dec 8 2015 00.00
Outliers	1

DATA QUALITY



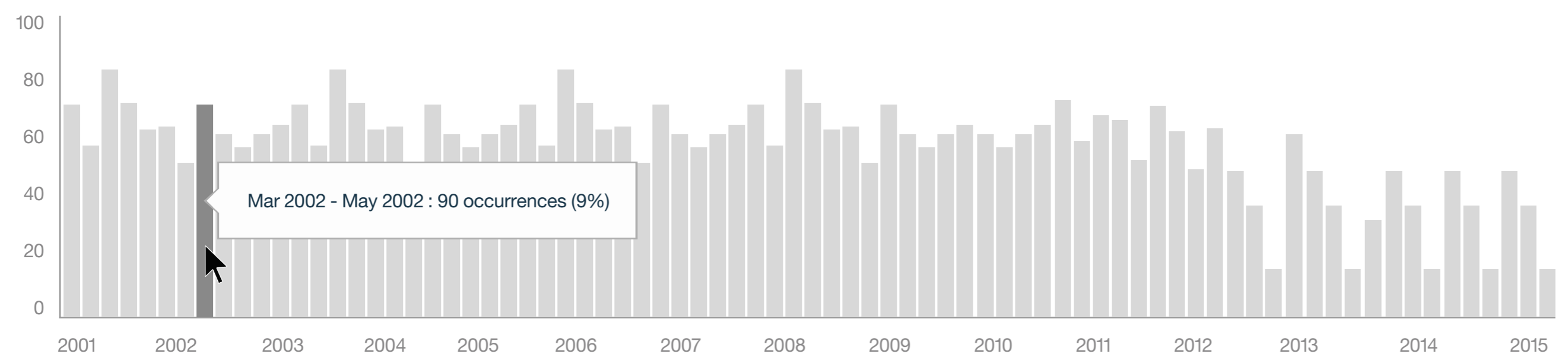
MISMATCHED VALUES AND OUTLIERS (258)

- 22/02/2015
- 1/9/2008
- 06/12/03
- 02-05-88
- 05-02-86

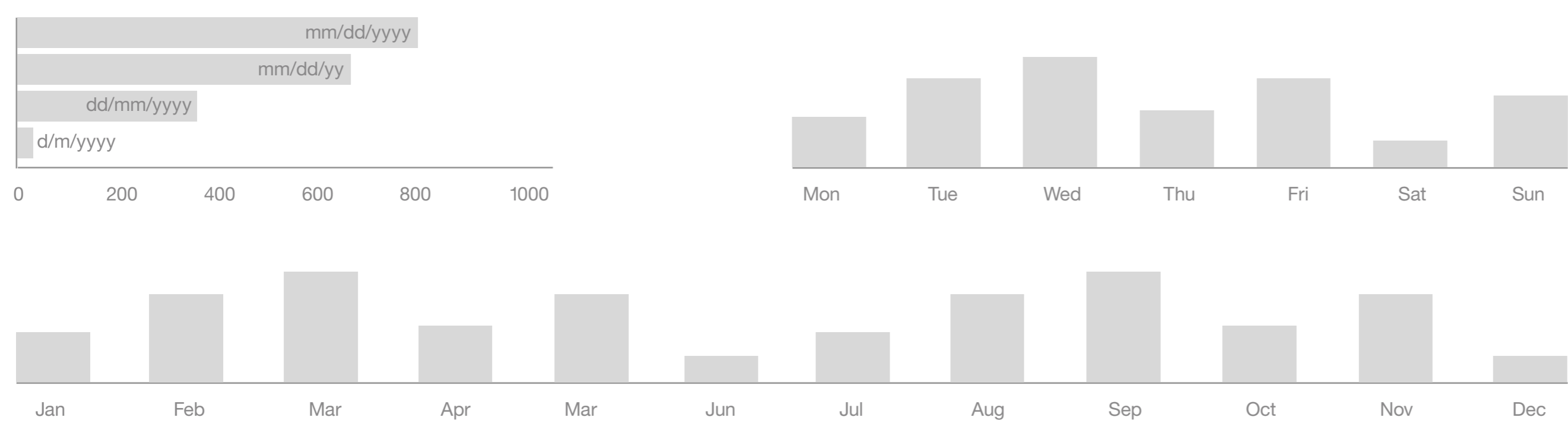
[View All](#)

FREQUENCY

↕ Bins : 75



PATTERNS







Select an operation or enter a command to cleanse and shape your data



Data Visualizations Metrics

Suggestions Data Flow

COLUMN NAME

Date



STATISTICS

Minimum	Jan 17 2001 00.00
Median	Aug 18 2008 00.00
Average	Jun 17 2008 06:25
Maximum	Dec 8 2015 00.00
Outliers	1

DATA QUALITY

Matched	75%	Mismatched	25%
Unique	100%		

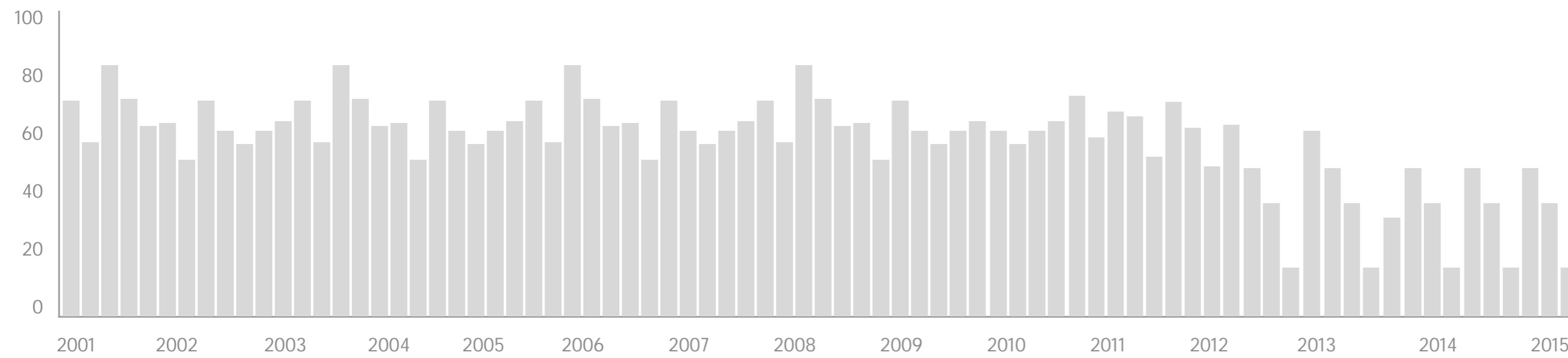
MISMATCHED VALUES AND OUTLIERS (258)

- 22/02/2015
- 1/9/2008
- 06/12/03
- 02-05-88
- 05-02-86

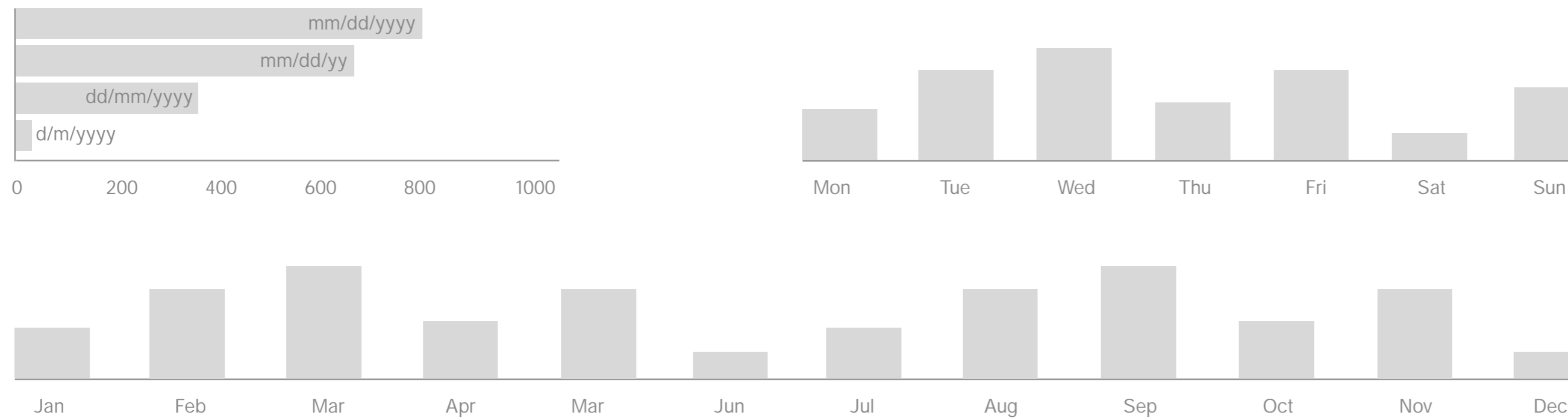
View All



FREQUENCY



PATTERNS





+ Select an operation or enter a command to cleanse and shape your data



COLUMN NAME Date

STATISTICS

Outliers	1
Minimum	Jan 17 2001 00:00
Median	Aug 18 2008 00:00
Maximum	Dec 8 2015 00:00
Average	Jun 17 2008 06:25

DATA QUALITY



INVALID AND OUTLIERS (258)

22/02/2015  
1/9/2008

MISMATCHED VALUES AND OUTLIERS

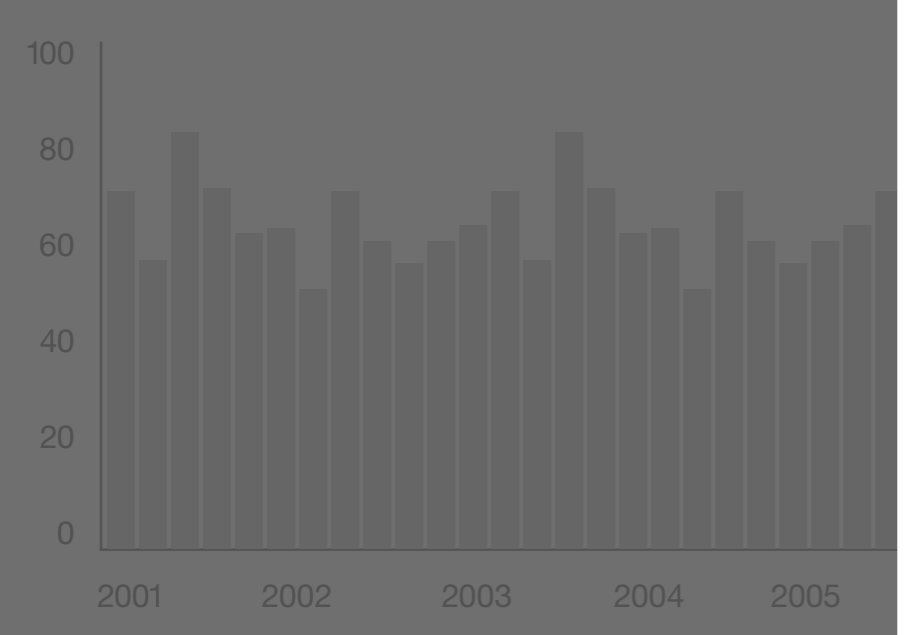
Mismatched (251)

- 22/02/2015
- 1/9/2008
- 06/12/03
- 02-05-88
- 05-02-86
- 22/02/2015
- 06/12/03
- 06/12/03
- 02-05-88
- 05-02-86
- 22/02/2015

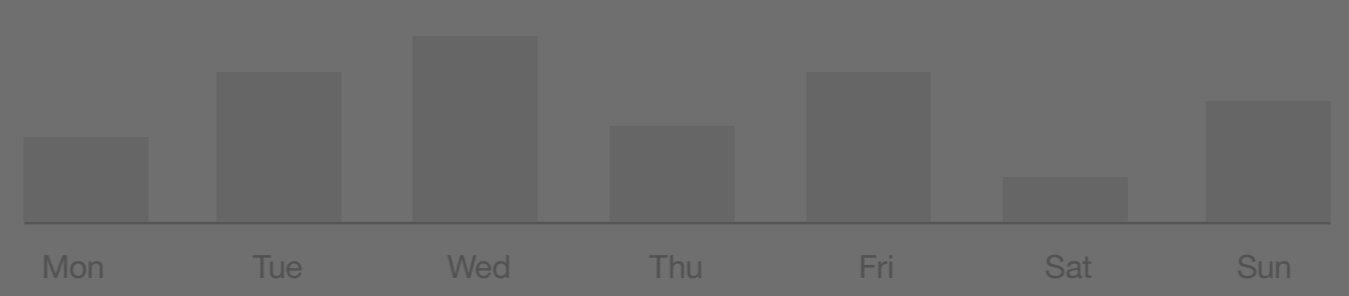
Outliers (7)

- 1/9/2008
- 6/2/3
- 02-05-888
- 005-02-86
- 22/02/2015
- 001/9/2008

FREQUENCY



PATTERNS



## Table Metrics (After MVP)



Data

Visualizations

Column metrics

Table metrics

Suggestions

Data Flow

78

Table-level quality

15%

Mismatched

75%

Missing

45%

Duplicates

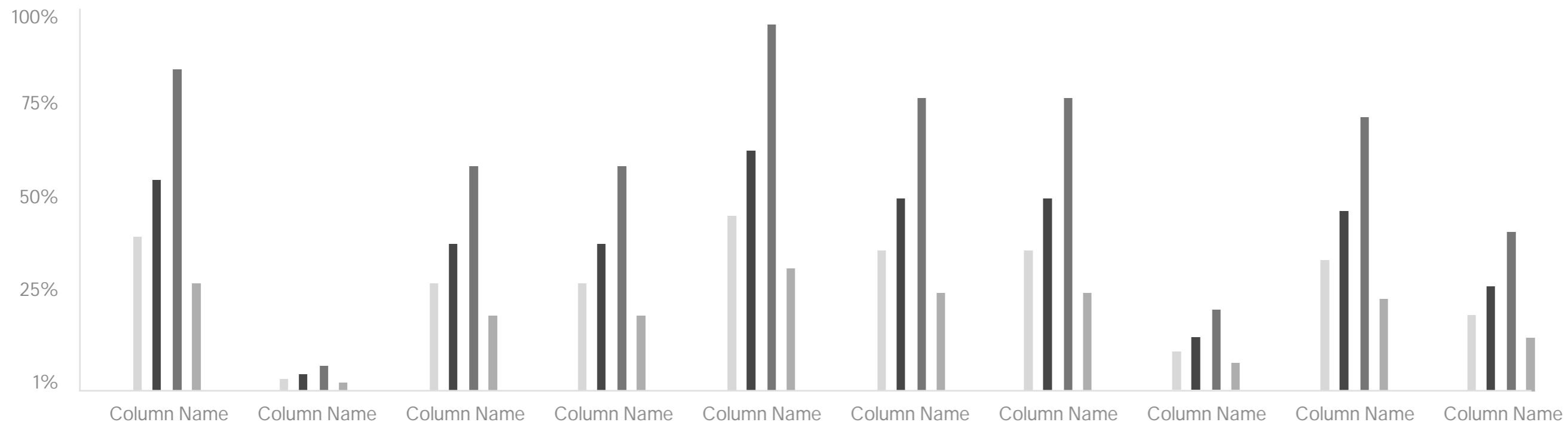
23%

Outliers

### Data quality problems

Distributed by column

% of all values in sample



#### TABLE METRICS

### Data quality

<Product Name> calculates data quality by using a sophisticated algorithm that rates data sets on a scale of 0 (poorest quality) to 100 (best quality). The algorithm searches for data quality problems at the column level, counting the occurrences and calculating the confidence of each problem that it identifies. This creates data quality scores for each column as well as a data quality score for the table (by averaging the column scores). The data quality problems that were found are listed below.

MISMATCHED

MISSING

DUPLICATES

OUTLIERS

