# Overview

Dataset statistics

|  |  |
| --- | --- |
| **Number of variables** | 19 |
| **Number of observations** | 507 |
| **Missing cells** | 60 |
| **Missing cells (%)** | 0.6% |
| **Duplicate rows** | 7 |
| **Duplicate rows (%)** | 1.4% |
| **Total size in memory** | 68.5 KiB |
| **Average record size in memory**  Variable types | 138.3 B |
| **Text** | 2 |
| **Boolean** | 3 |
| **Categorical** | 8 |
| **DateTime** | 1 |

**Numeric** 5

Alerts

Dataset has 7 (1.4%)

duplicate rows

**Duplicates**

customer\_satisfaction

is highly overall correlated with

is\_expedited\_delivery

**High correlation**

delivery\_charges

is highly overall correlated with

is\_expedited\_delivery

**High correlation**

education

is highly overall correlated with

occupation

**High correlation**

is\_expedited\_delivery

is highly overall correlated with

customer\_satisfaction

and

other

1

fi

elds

(customer\_satisfaction,

delivery\_charges)

**High correlation**

occupation

is highly overall correlated with

education

**High correlation**

order\_payment

is highly overall correlated with

order\_type

**High correlation**

order\_price

is highly overall correlated with

total\_bill

**High correlation**

order\_type

is highly overall correlated with

order\_payment

**High correlation**

total\_bill

is highly overall correlated with

order\_price

**High correlation**

warehouse\_area\_type

is highly overall correlated with

warehouse\_name

and

1

other

fi

elds

(warehouse\_name, warehouse\_size)

**High correlation**

warehouse\_name

is highly overall correlated with

warehouse\_area\_type

and

1

other

fi

elds

(warehouse\_area\_type, warehouse\_size

)

**High correlation**

warehouse\_size

is highly overall correlated with

warehouse\_area\_type

and

1

other

fi

elds

(warehouse\_area\_type, warehouse\_name)

**High correlation**

education

has 10 (2.0%) missing values

**Missing**

delivery\_charges

has 10 (2.0%) missing values

**Missing**

order\_type

has 23 (4.5%) missing values

**Missing**

order\_payment

has 17 (3.4%) missing values

**Missing**

Reproduction

|  |  |
| --- | --- |
| **Analysis started** | 2024-04-05 19:37:16.730432 |
| **Analysis finished** | 2024-04-05 19:37:26.177726 |
| **Duration** | 9.45 seconds |
| **Software version** | [ydata-profiling vv4.7.0 (https://github.com/ydataai/ydata-profiling)](https://github.com/ydataai/ydata-profiling) |

**Download** config.json (data:text/plain;charset=utf-

**configuration** 8,%7B%22title%22%3A%20%22Pandas%20Profiling%20Report%22%2C%20%22dataset%22%3A%20%7B%22description%22%3A%20%22%

# Variables

|  |
| --- |
| Select Columns |

customer\_id

Text

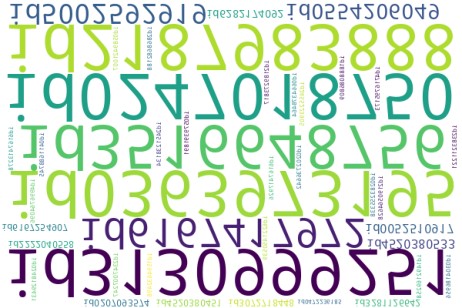
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Distinct** | |  | | 500 |  | |
| **Distinct (%)** | |  | | 98.6% |  | |
| **Missing** | |  | | 0 |  | |
| **Missing (%)** | |  | | 0.0% |  | |
| **Memory size**  Length | |  | | 4.1 KiB |  | |
| **Max length** | |  | |  | 12 | |
| **Median length** | |  | |  | 12 | |
| **Mean length** | |  | |  | 12 | |
| **Min length**  Characters and Unicode | |  | |  | 12 | |
| **Total characters** | | 6084 | |  |  | |
| **Distinct characters** | | 12 | |  |  | |
| **Distinct categories** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_character\_property#General\_Category)](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) | | | [**?**](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) | |
| **Distinct scripts** | | 1 [(https://en.wikipedia.org/wiki/Script\_(Unicode)#List\_of\_scripts\_in\_Unicode)](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) | | | [**?**](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) | |
| **Distinct blocks** | 1 [(https://en.wikipedia.org/wiki/Unicode\_block)](https://en.wikipedia.org/wiki/Unicode_block) | | | | | [**?**](https://en.wikipedia.org/wiki/Unicode_block) |
| Unique | The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables. | | | | |  |
| **Unique** |  | | 493 | | | **?** |
| **Unique (%)**  Sample |  | | 97.2% | | |  |
| **1st row** | ID0571731474 | |  | | |  |
| **2nd row** | ID6167441063 | |  | | |  |
| **3rd row** | ID6221450539 | |  | | |  |
| **4th row** | ID2383211221 | |  | | |  |

**5th row** ID0248129431

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| id0363973195 | 2 | 0.4% |
| id0247018750 | 2 | 0.4% |
| id3130999251 | 2 | 0.4% |
| id2187983888 | 2 | 0.4% |
| id3516648756 | 2 | 0.4% |
| id6167417972 | 2 | 0.4% |
| id5002592919 | 2 | 0.4% |
| id0554206049 | 1 | 0.2% |
| id0052510917 | 1 | 0.2% |
| id4520380533 | 1 | 0.2% |

Other values (490) 490 96.6%



Most occurring characters

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 0 | 686 | 11.3% |
| 1 | 559 | 9.2% |
| 4 | 543 | 8.9% |
| 2 | 526 | 8.6% |
| I | 507 | 8.3% |
| D | 507 | 8.3% |
| 7 | 505 | 8.3% |
| 6 | 490 | 8.1% |
| 3 | 483 | 7.9% |
| 5 | 460 | 7.6% |

Other values (2) 818 13.4%

Most occurring categories

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 6084 | 100.0% |
| Most frequent character per category  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| 0 | 686 | 11.3% |
| 1 | 559 | 9.2% |
| 4 | 543 | 8.9% |
| 2 | 526 | 8.6% |
| I | 507 | 8.3% |
| D | 507 | 8.3% |
| 7 | 505 | 8.3% |
| 6 | 490 | 8.1% |
| 3 | 483 | 7.9% |
| 5 | 460 | 7.6% |

Other values (2) 818 13.4%

Most occurring scripts

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 6084 | 100.0% |
| Most frequent character per script  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| 0 | 686 | 11.3% |
| 1 | 559 | 9.2% |
| 4 | 543 | 8.9% |
| 2 | 526 | 8.6% |
| I | 507 | 8.3% |
| D | 507 | 8.3% |
| 7 | 505 | 8.3% |
| 6 | 490 | 8.1% |
| 3 | 483 | 7.9% |
| 5 | 460 | 7.6% |

Other values (2) 818 13.4%

Most occurring blocks

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 6084 | 100.0% |
| Most frequent character per block  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| 0 | 686 | 11.3% |
| 1 | 559 | 9.2% |
| 4 | 543 | 8.9% |
| 2 | 526 | 8.6% |
| I | 507 | 8.3% |
| D | 507 | 8.3% |
| 7 | 505 | 8.3% |
| 6 | 490 | 8.1% |
| 3 | 483 | 7.9% |
| 5 | 460 | 7.6% |

Other values (2) 818 13.4%

marital\_status

Boolean

|  |  |  |
| --- | --- | --- |
| **Distinct** | 2 |  |
| **Distinct (%)** | 0.4% |  |
| **Missing** | 0 |  |
| **Missing (%)** | 0.0% |  |
| **Memory size** | 635.0 B |  |
| **Value** | **Count** | **Frequency (%)** |
| True | 272 | 53.6% |

False 235 46.4%

|  |  |
| --- | --- |
|  |  |

occupation

Categorical

## HIGH CORRELATION (This variable has a high overall correlation with 1 fields: education)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Distinct** | |  | 9 |  |
| **Distinct (%)** | |  | 1.8% |  |
| **Missing** | |  | 0 |  |
| **Missing (%)** | |  | 0.0% |  |
| **Memory size**  Length | |  | 4.1 KiB |  |
| **Max length** | |  | 12 |  |
| **Median length** | |  | 10 |  |
| **Mean length** | |  | 8.8126233 |  |
| **Min length**  Characters and Unicode | |  | 6 |  |
| **Total characters** | | 4468 |  |  |
| **Distinct characters** | | 28 |  |  |
| **Distinct categories** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_character\_property#General\_Category)](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) | | [**?**](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) |
| **Distinct scripts** | | 1 [(https://en.wikipedia.org/wiki/Script\_(Unicode)#List\_of\_scripts\_in\_Unicode)](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) | | [**?**](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) |
| **Distinct blocks** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_block)](https://en.wikipedia.org/wiki/Unicode_block) | | [**?**](https://en.wikipedia.org/wiki/Unicode_block) |
| Unique | The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables. | | |  |
| **Unique** | 0 | | | **?** |
| **Unique (%)**  Sample | 0.0% | | |  |
| **1st row** | Manager | | |  |
| **2nd row** | Clerical | | |  |
| **3rd row** | Lawyer | | |  |
| **4th row** | Manager | | |  |

**5th row** Student

Common Values

**Value Count Frequency (%)**

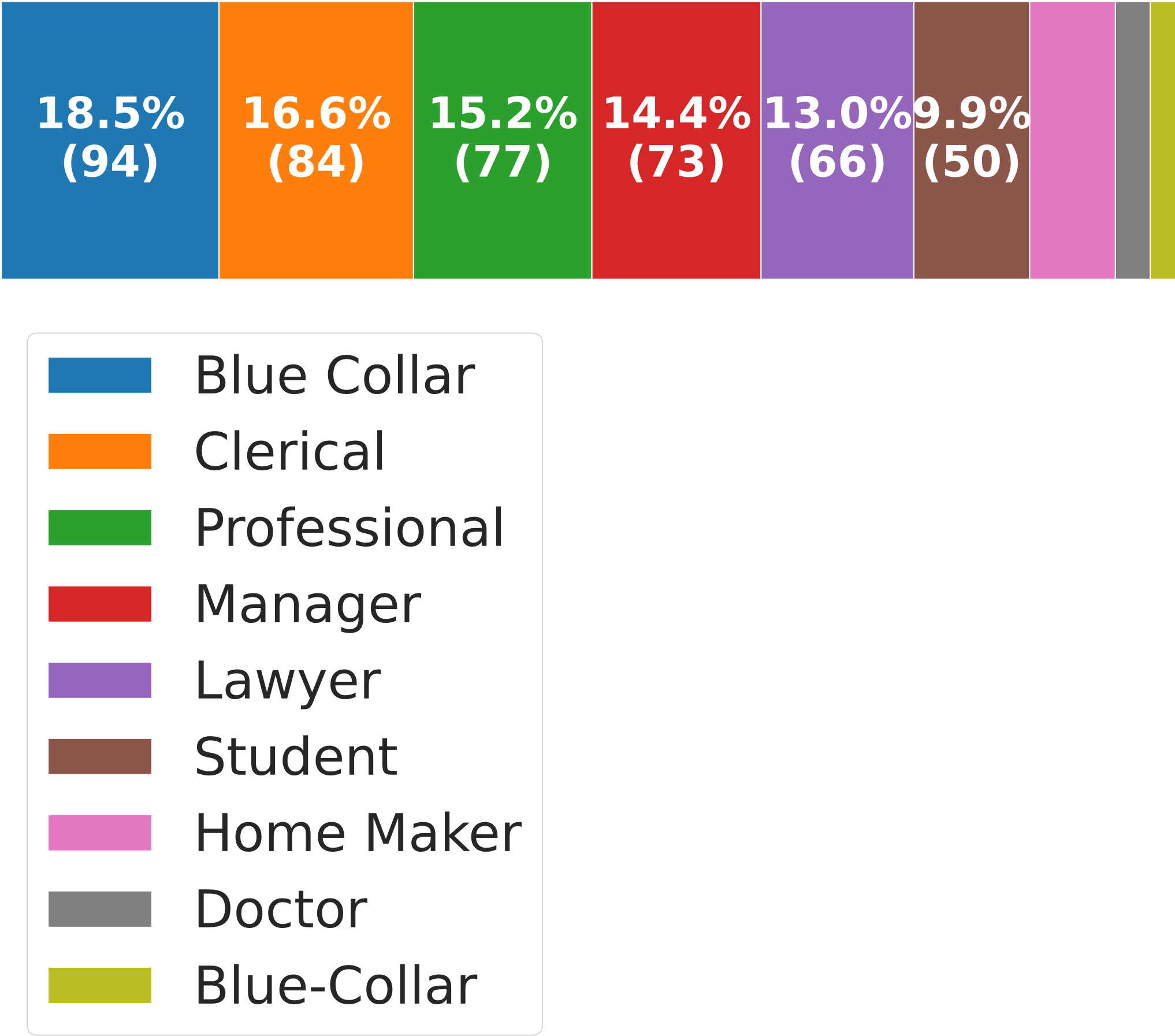
|  |  |  |
| --- | --- | --- |
| Blue Collar | 94 | 18.5% |
| Clerical | 84 | 16.6% |
| Professional | 77 | 15.2% |
| Manager | 73 | 14.4% |
| Lawyer | 66 | 13.0% |
| Student | 50 | 9.9% |
| Home Maker | 37 | 7.3% |
| Doctor | 15 | 3.0% |

Blue-Collar 11 2.2%

Length

Histogram of lengths of the category

Common Values (Plot)



**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| blue | 94 | 14.7% |
| collar | 94 | 14.7% |
| clerical | 84 | 13.2% |
| professional | 77 | 12.1% |
| manager | 73 | 11.4% |
| lawyer | 66 | 10.3% |
| student | 50 | 7.8% |
| home | 37 | 5.8% |
| maker | 37 | 5.8% |

doctor 15 2.4%

Most occurring characters

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| l | 560 | 12.5% |
| e | 529 | 11.8% |
| a | 515 | 11.5% |
| r | 457 | 10.2% |
| o | 326 | 7.3% |
| n | 200 | 4.5% |
| C | 189 | 4.2% |
| i | 161 | 3.6% |
| u | 155 | 3.5% |
| s | 154 | 3.4% |

Other values (18) 1222 27.4%

Most occurring categories

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4468 | 100.0% |
| Most frequent character per category  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| l | 560 | 12.5% |
| e | 529 | 11.8% |
| a | 515 | 11.5% |
| r | 457 | 10.2% |
| o | 326 | 7.3% |
| n | 200 | 4.5% |
| C | 189 | 4.2% |
| i | 161 | 3.6% |
| u | 155 | 3.5% |
| s | 154 | 3.4% |

Other values (18) 1222 27.4%

Most occurring scripts

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4468 | 100.0% |
| Most frequent character per script  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| l | 560 | 12.5% |
| e | 529 | 11.8% |
| a | 515 | 11.5% |
| r | 457 | 10.2% |
| o | 326 | 7.3% |
| n | 200 | 4.5% |
| C | 189 | 4.2% |
| i | 161 | 3.6% |
| u | 155 | 3.5% |
| s | 154 | 3.4% |

Other values (18) 1222 27.4%

Most occurring blocks

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4468 | 100.0% |
| Most frequent character per block  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| l | 560 | 12.5% |
| e | 529 | 11.8% |
| a | 515 | 11.5% |
| r | 457 | 10.2% |
| o | 326 | 7.3% |
| n | 200 | 4.5% |
| C | 189 | 4.2% |
| i | 161 | 3.6% |
| u | 155 | 3.5% |
| s | 154 | 3.4% |

Other values (18) 1222 27.4%

education

Categorical

## HIGH CORRELATION (This variable has a high overall correlation with 1 fields: occupation) MISSING

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Distinct** | |  | 6 |  |
| **Distinct (%)** | |  | 1.2% |  |
| **Missing** | |  | 10 |  |
| **Missing (%)** | |  | 2.0% |  |
| **Memory size**  Length | |  | 4.1 KiB |  |
| **Max length** | |  | 11 |  |
| **Median length** | |  | 9 |  |
| **Mean length** | |  | 8.1448692 |  |
| **Min length**  Characters and Unicode | |  | 1 |  |
| **Total characters** | | 4048 |  |  |
| **Distinct characters** | | 19 |  |  |
| **Distinct categories** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_character\_property#General\_Category)](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) | | [**?**](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) |
| **Distinct scripts** | | 1 [(https://en.wikipedia.org/wiki/Script\_(Unicode)#List\_of\_scripts\_in\_Unicode)](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) | | [**?**](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) |
| **Distinct blocks** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_block)](https://en.wikipedia.org/wiki/Unicode_block) | | [**?**](https://en.wikipedia.org/wiki/Unicode_block) |
| Unique | The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables. | | |  |
| **Unique** | 0 | | | **?** |
| **Unique (%)**  Sample | 0.0% | | |  |
| **1st row** | Masters | | |  |
| **2nd row** | School | | |  |
| **3rd row** | Masters | | |  |
| **4th row** | High School | | |  |

**5th row** School

Common Values

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| High School | 146 | 28.8% |
| Bachelors | 137 | 27.0% |
| Masters | 97 | 19.1% |
| School | 69 | 13.6% |
| PhD | 34 | 6.7% |
| ? | 14 | 2.8% |

(Missing) 10 2.0%

Length

Histogram of lengths of the category

Common Values (Plot)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| school | 215 | 33.4% |
| high | 146 | 22.7% |
| bachelors | 137 | 21.3% |
| masters | 97 | 15.1% |
| phd | 34 | 5.3% |

14 2.2%

Most occurring characters

**Value Count Frequency (%)**

|  |  |  |  |
| --- | --- | --- | --- |
| o | 567 | 14.0% | |
| h | 532 | 13.1% | |
| c | 352 | 8.7% |  |
| l | 352 | 8.7% |  |
| s | 331 |  | 8.2% |
| a | 234 |  | 5.8% |
| e | 234 |  | 5.8% |
| r | 234 |  | 5.8% |
| S | 215 |  | 5.3% |
| H | 146 |  | 3.6% |

Other values (9) 851 21.0%

Most occurring categories

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4048 | 100.0% |
| Most frequent character per category  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| o | 567 | 14.0% |
| h | 532 | 13.1% |
| c | 352 | 8.7% |
| l | 352 | 8.7% |
| s | 331 | 8.2% |
| a | 234 | 5.8% |
| e | 234 | 5.8% |
| r | 234 | 5.8% |
| S | 215 | 5.3% |
| H | 146 | 3.6% |

Other values (9) 851 21.0%

Most occurring scripts

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4048 | 100.0% |
| Most frequent character per script  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| o | 567 | 14.0% |
| h | 532 | 13.1% |
| c | 352 | 8.7% |
| l | 352 | 8.7% |
| s | 331 | 8.2% |
| a | 234 | 5.8% |
| e | 234 | 5.8% |
| r | 234 | 5.8% |
| S | 215 | 5.3% |
| H | 146 | 3.6% |

Other values (9) 851 21.0%

Most occurring blocks

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4048 | 100.0% |
| Most frequent character per block  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| o | 567 | 14.0% |
| h | 532 | 13.1% |
| c | 352 | 8.7% |
| l | 352 | 8.7% |
| s | 331 | 8.2% |
| a | 234 | 5.8% |
| e | 234 | 5.8% |
| r | 234 | 5.8% |
| S | 215 | 5.3% |
| H | 146 | 3.6% |

Other values (9) 851 21.0%

warehouse\_name

Categorical

## HIGH CORRELATION (This variable has a high overall correlation with 2 fields: warehouse\_area\_type, warehouse\_size)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Distinct** | |  | 3 |  |
| **Distinct (%)** | |  | 0.6% |  |
| **Missing** | |  | 0 |  |
| **Missing (%)** | |  | 0.0% |  |
| **Memory size**  Length | |  | 4.1 KiB |  |
| **Max length** | |  | 9 |  |
| **Median length** | |  | 8 |  |
| **Mean length** | |  | 7.8954635 |  |
| **Min length**  Characters and Unicode | |  | 6 |  |
| **Total characters** | | 4003 |  |  |
| **Distinct characters** | | 16 |  |  |
| **Distinct categories** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_character\_property#General\_Category)](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) | | [**?**](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) |
| **Distinct scripts** | | 1 [(https://en.wikipedia.org/wiki/Script\_(Unicode)#List\_of\_scripts\_in\_Unicode)](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) | | [**?**](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) |
| **Distinct blocks** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_block)](https://en.wikipedia.org/wiki/Unicode_block) | | [**?**](https://en.wikipedia.org/wiki/Unicode_block) |
| Unique | The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables. | | |  |
| **Unique** | 0 | | | **?** |
| **Unique (%)**  Sample | 0.0% | | |  |
| **1st row** | Nickolson | | |  |
| **2nd row** | Thompson | | |  |
| **3rd row** | Nickolson | | |  |
| **4th row** | Nickolson | | |  |

**5th row** Nickolson

Common Values

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| Thompson | 200 | 39.4% |
| Nickolson | 187 | 36.9% |

Bakers 120 23.7%

Length

Histogram of lengths of the category

Common Values (Plot)

|  |  |  |
| --- | --- | --- |
|  |  |  |

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| thompson | 200 | 39.4% |
| nickolson | 187 | 36.9% |

bakers 120 23.7%

Most occurring characters

**Value Count Frequency (%)**

|  |  |  |  |
| --- | --- | --- | --- |
| o | 774 | 19.3% | |
| s | 507 | 12.7% | |
| n | 387 | 9.7% |  |
| k | 307 |  | 7.7% |
| T | 200 |  | 5.0% |
| h | 200 |  | 5.0% |
| m | 200 |  | 5.0% |
| p | 200 |  | 5.0% |
| N | 187 |  | 4.7% |
| i | 187 |  | 4.7% |

Other values (6) 854 21.3%

Most occurring categories

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4003 | 100.0% |
| Most frequent character per category  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| o | 774 | 19.3% |
| s | 507 | 12.7% |
| n | 387 | 9.7% |
| k | 307 | 7.7% |
| T | 200 | 5.0% |
| h | 200 | 5.0% |
| m | 200 | 5.0% |
| p | 200 | 5.0% |
| N | 187 | 4.7% |
| i | 187 | 4.7% |

Other values (6) 854 21.3%

Most occurring scripts

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4003 | 100.0% |
| Most frequent character per script  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| o | 774 | 19.3% |
| s | 507 | 12.7% |
| n | 387 | 9.7% |
| k | 307 | 7.7% |
| T | 200 | 5.0% |
| h | 200 | 5.0% |
| m | 200 | 5.0% |
| p | 200 | 5.0% |
| N | 187 | 4.7% |
| i | 187 | 4.7% |

Other values (6) 854 21.3%

Most occurring blocks

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4003 | 100.0% |
| Most frequent character per block  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| o | 774 | 19.3% |
| s | 507 | 12.7% |
| n | 387 | 9.7% |
| k | 307 | 7.7% |
| T | 200 | 5.0% |
| h | 200 | 5.0% |
| m | 200 | 5.0% |
| p | 200 | 5.0% |
| N | 187 | 4.7% |
| i | 187 | 4.7% |

Other values (6) 854 21.3%

order\_id

0.0%

Text

**Distinct**

500

**Distinct (%)**

98.6

%

**Missing**

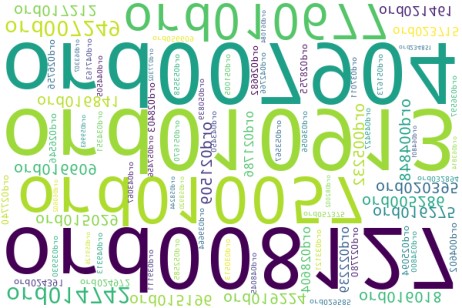
0

**Missing (%)**

**Memory size**

4.1

KiB



Length

**Max length**

9

**Median length**

9

**Mean length**

9

**Min length**

9

Characters and Unicode

[**?**](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)

[**?**](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)

[**?**](https://en.wikipedia.org/wiki/Unicode_block)

**Total characters**

4563

**Distinct characters**

13

**Distinct categories**

1

[(](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)

<https://en.wikipedia.org/wiki/Unicode_character_property#General_Category>

[)](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)

**Distinct scripts**

1

[(](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)

<https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode>

[)](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)

**Distinct blocks**

1

[(](https://en.wikipedia.org/wiki/Unicode_block)

<https://en.wikipedia.org/wiki/Unicode_block>

[)](https://en.wikipedia.org/wiki/Unicode_block)

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

Unique

**Unique** 493

**?**

**Unique (%)** 97.2%

Sample

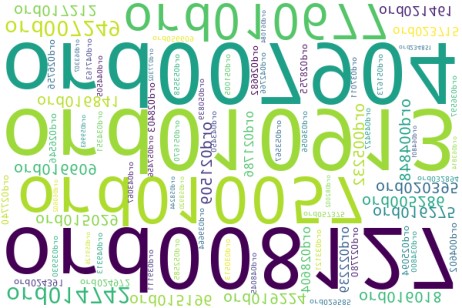
|  |  |
| --- | --- |
| **1st row** | ORD001041 |
| **2nd row** | ORD004361 |
| **3rd row** | ORD004602 |
| **4th row** | ORD004848 |

**5th row** ORD005232

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| ord010913 | 2 | 0.4% |
| ord007904 | 2 | 0.4% |
| ord008127 | 2 | 0.4% |
| ord010057 | 2 | 0.4% |
| ord010677 | 2 | 0.4% |
| ord014742 | 2 | 0.4% |
| ord007249 | 2 | 0.4% |
| ord016275 | 1 | 0.2% |
| ord020395 | 1 | 0.2% |
| ord019224 | 1 | 0.2% |

Other values (490) 490 96.6%



Most occurring characters

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| O | 507 | 11.1% |
| R | 507 | 11.1% |
| D | 507 | 11.1% |
| 0 | 404 | 8.9% |
| 2 | 363 | 8.0% |
| 3 | 358 | 7.8% |
| 1 | 355 | 7.8% |
| 4 | 341 | 7.5% |
| 5 | 266 | 5.8% |
| 7 | 260 | 5.7% |

Other values (3) 695 15.2%

Most occurring categories

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4563 | 100.0% |
| Most frequent character per category  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| O | 507 | 11.1% |
| R | 507 | 11.1% |
| D | 507 | 11.1% |
| 0 | 404 | 8.9% |
| 2 | 363 | 8.0% |
| 3 | 358 | 7.8% |
| 1 | 355 | 7.8% |
| 4 | 341 | 7.5% |
| 5 | 266 | 5.8% |
| 7 | 260 | 5.7% |

Other values (3) 695 15.2%

Most occurring scripts

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4563 | 100.0% |
| Most frequent character per script  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| O | 507 | 11.1% |
| R | 507 | 11.1% |
| D | 507 | 11.1% |
| 0 | 404 | 8.9% |
| 2 | 363 | 8.0% |
| 3 | 358 | 7.8% |
| 1 | 355 | 7.8% |
| 4 | 341 | 7.5% |
| 5 | 266 | 5.8% |
| 7 | 260 | 5.7% |

Other values (3) 695 15.2%

Most occurring blocks

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4563 | 100.0% |
| Most frequent character per block  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| O | 507 | 11.1% |
| R | 507 | 11.1% |
| D | 507 | 11.1% |
| 0 | 404 | 8.9% |
| 2 | 363 | 8.0% |
| 3 | 358 | 7.8% |
| 1 | 355 | 7.8% |
| 4 | 341 | 7.5% |
| 5 | 266 | 5.8% |
| 7 | 260 | 5.7% |

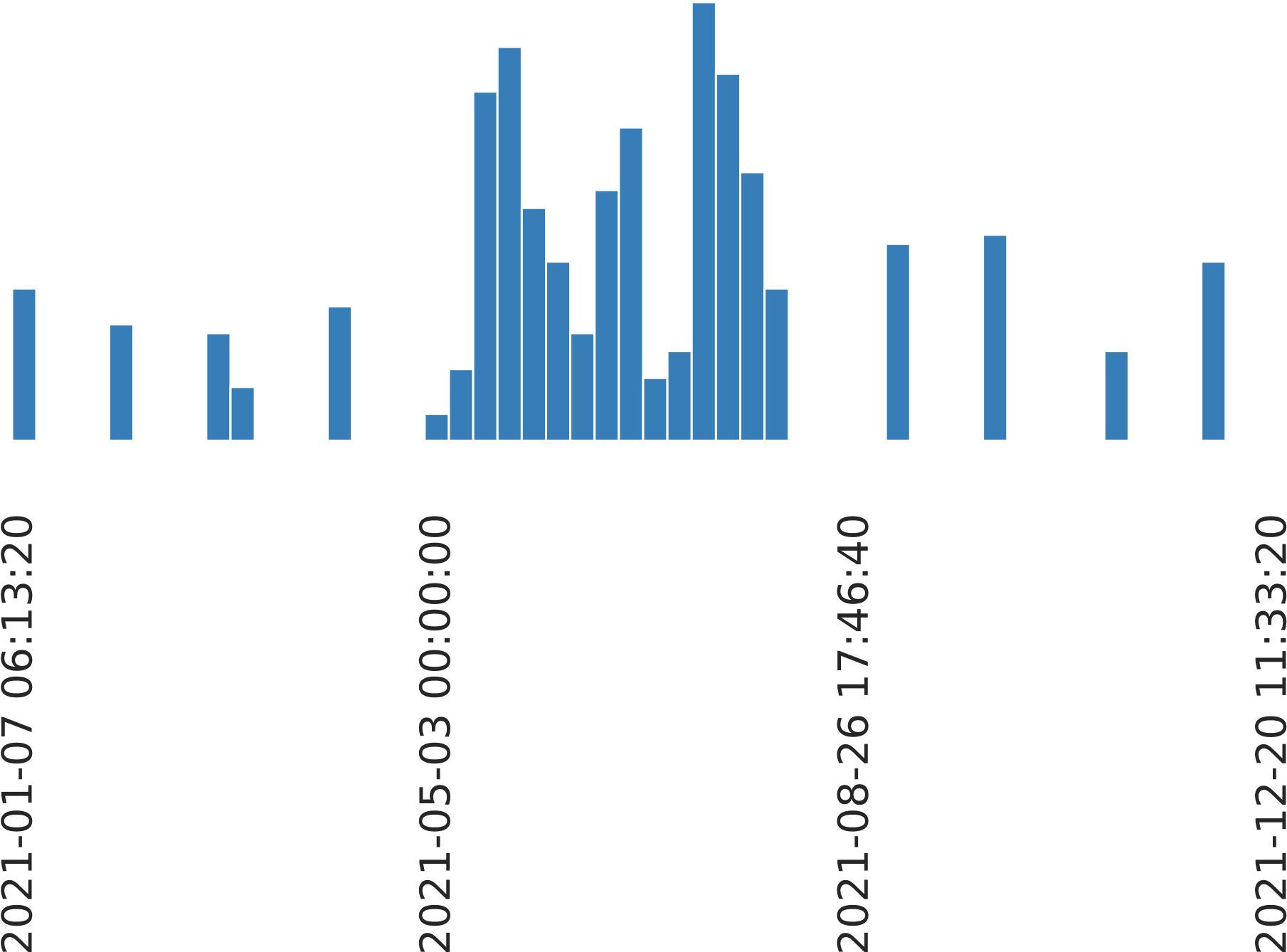
Other values (3) 695 15.2%

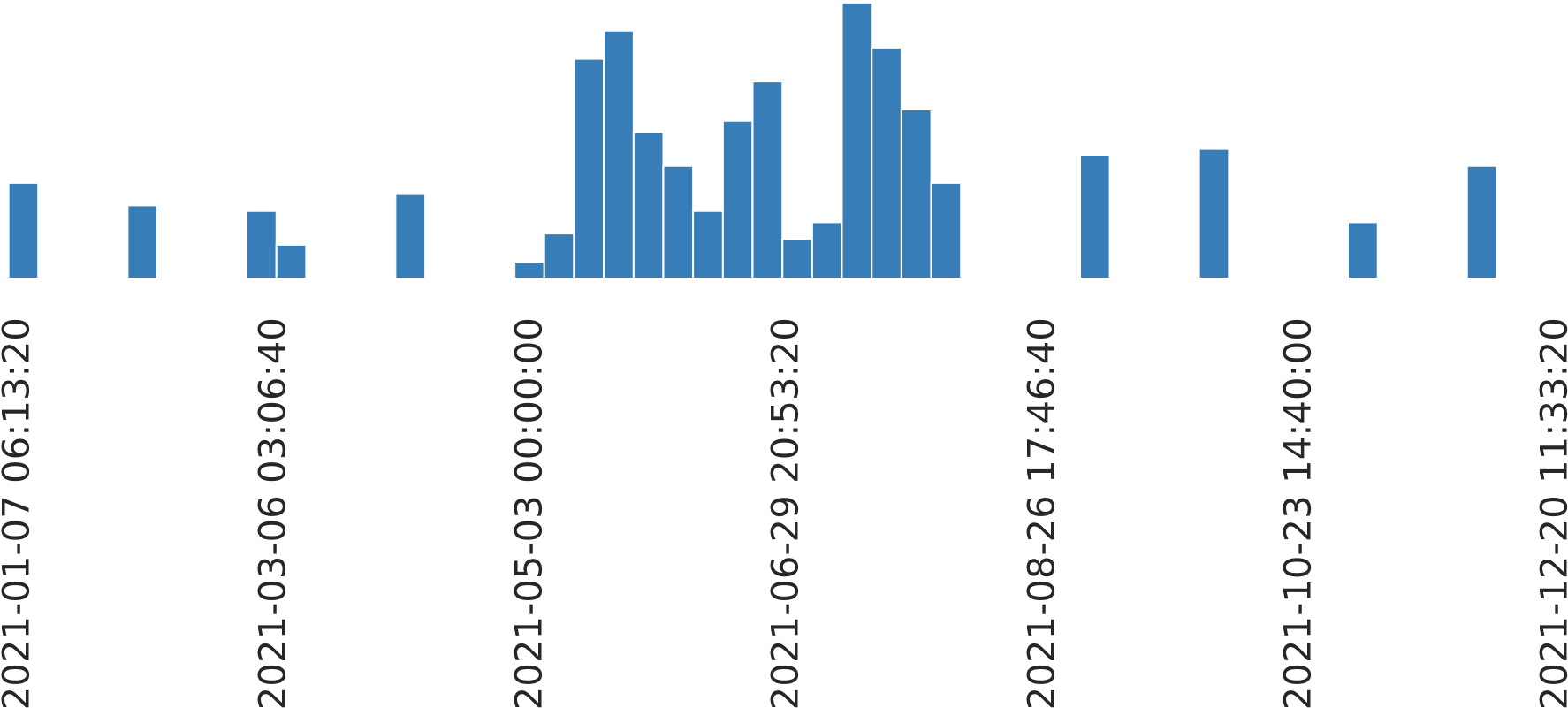
order\_date

Date

|  |  |  |
| --- | --- | --- |
| **Distinct** |  | 91 |
| **Distinct (%)** |  | 17.9% |
| **Missing** |  | 0 |
| **Missing (%)** |  | 0.0% |
| **Memory size** |  | 4.1 KiB |
| **Minimum** | 2021-01-05 00:00:00 |  |

**Maximum** 2021-12-07 00:00:00





**Histogram with fixed size bins** (bins=50)

order\_price

Real number (ℝ)

## HIGH CORRELATION (This variable has a high overall correlation with 1 fields: total\_bill)

|  |  |
| --- | --- |
| **Distinct** | 414 |
| **Distinct (%)** | 81.7% |
| **Missing** | 0 |
| **Missing (%)** | 0.0% |
| **Infinite** | 0 |
| **Infinite (%)** | 0.0% |
| **Mean** | 25169.74 |
| **Minimum** | 0 |
| **Maximum** | 947691 |
| **Zeros** | 5 |
| **Zeros (%)** | 1.0% |
| **Negative** | 0 |
| **Negative (%)** | 0.0% |

|  |  |
| --- | --- |
| **Minimum** | 0 |
| **5-th percentile** | 2493.5 |
| **Q1** | 6877.5 |

**Memory size** 4.1 KiB

Quantile statistics

|  |  |
| --- | --- |
| **median** | 12765 |
| **Q3** | 20325 |
| **95-th percentile** | 31898.5 |
| **Maximum** | 947691 |
| **Range** | 947691 |
| **Interquartile range (IQR)** Descriptive statistics | 13447.5 |
| **Standard deviation** | 85785.462 |
| **Coefficient of variation (CV)** | 3.4082777 |
| **Kurtosis** | 67.457024 |
| **Mean** | 25169.74 |
| **Median Absolute Deviation (MAD)** | 6305 |
| **Skewness** | 8.0023731 |
| **Sum** | 12761058 |
| **Variance** | 7.3591456 × 109 |

**Monotonicity** Not monotonic

**Histogram with fixed size bins** (bins=50)

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 0 | 5 | 1.0% |
| 7760 | 4 | 0.8% |
| 5910 | 3 | 0.6% |
| 11100 | 3 | 0.6% |
| 6910 | 3 | 0.6% |
| 8770 | 3 | 0.6% |
| 7050 | 3 | 0.6% |
| 6630 | 3 | 0.6% |
| 8995 | 3 | 0.6% |
| 4405 | 3 | 0.6% |

Other values (404) 474 93.5%

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 0 | 5 | 1.0% |
| 585 | 1 | 0.2% |
| 730 | 2 | 0.4% |
| 1010 | 3 | 0.6% |
| 1380 | 1 | 0.2% |
| 1525 | 1 | 0.2% |
| 1619 | 1 | 0.2% |
| 1655 | 1 | 0.2% |
| 1660 | 2 | 0.4% |

2085 2 0.4%

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 947691 | 1 | 0.2% |
| 865916 | 1 | 0.2% |
| 672832 | 1 | 0.2% |
| 655164 | 1 | 0.2% |
| 653959 | 1 | 0.2% |
| 597240 | 1 | 0.2% |
| 461736 | 1 | 0.2% |
| 454996 | 1 | 0.2% |
| 386776 | 1 | 0.2% |

98614 1 0.2%

delivery\_charges

Real number (ℝ)

## HIGH CORRELATION (This variable has a high overall correlation with 1 fields: is\_expedited\_delivery) MISSING

|  |  |
| --- | --- |
| **Distinct** | 456 |
| **Distinct (%)** | 91.8% |
| **Missing** | 10 |
| **Missing (%)** | 2.0% |
| **Infinite** | 0 |
| **Infinite (%)** | 0.0% |
| **Mean** | 76.099698 |
| **Minimum** | 0 |
| **Maximum** | 114.04 |
| **Zeros** | 4 |
| **Zeros (%)** | 0.8% |
| **Negative** | 0 |
| **Negative (%)** | 0.0% |

|  |  |
| --- | --- |
| **Minimum** | 0 |
| **5-th percentile** | 52.864 |
| **Q1** | 65.87 |

**Memory size** 4.1 KiB

Quantile statistics

|  |  |
| --- | --- |
| **median** | 76.29 |
| **Q3** | 82.44 |
| **95-th percentile** | 105.052 |
| **Maximum** | 114.04 |
| **Range** | 114.04 |
| **Interquartile range (IQR)** Descriptive statistics | 16.57 |
| **Standard deviation** | 16.026792 |
| **Coefficient of variation (CV)** | 0.21060257 |
| **Kurtosis** | 3.0859553 |
| **Mean** | 76.099698 |
| **Median Absolute Deviation (MAD)** | 9.77 |
| **Skewness** | -0.40788841 |
| **Sum** | 37821.55 |
| **Variance** | 256.85806 |

**Monotonicity** Not monotonic

**Histogram with fixed size bins** (bins=50)

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 0 | 4 | 0.8% |
| 108.58 | 3 | 0.6% |
| 79.92 | 2 | 0.4% |
| 74.21 | 2 | 0.4% |
| 75.54 | 2 | 0.4% |
| 85.48 | 2 | 0.4% |
| 62.82 | 2 | 0.4% |
| 65.5 | 2 | 0.4% |
| 64.81 | 2 | 0.4% |
| 82.37 | 2 | 0.4% |
| Other values (446) | 474 | 93.5% |

(Missing) 10 2.0%

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 0 | 4 | 0.8% |
| 46.35 | 1 | 0.2% |
| 46.73 | 1 | 0.2% |
| 46.89 | 1 | 0.2% |
| 48.55 | 1 | 0.2% |
| 48.71 | 1 | 0.2% |
| 48.99 | 1 | 0.2% |
| 49.88 | 1 | 0.2% |
| 50.05 | 1 | 0.2% |

50.65 1 0.2%

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 114.04 | 1 | 0.2% |
| 113.83 | 1 | 0.2% |
| 112.42 | 1 | 0.2% |
| 112.04 | 1 | 0.2% |
| 110.36 | 1 | 0.2% |
| 110.2 | 1 | 0.2% |
| 110.17 | 1 | 0.2% |
| 108.92 | 1 | 0.2% |
| 108.87 | 1 | 0.2% |

108.58 3 0.6%

coupon\_discount

Categorical

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Distinct** | |  | 5 |  |
| **Distinct (%)** | |  | 1.0% |  |
| **Missing** | |  | 0 |  |
| **Missing (%)** | |  | 0.0% |  |
| **Memory size**  Length | |  | 4.1 KiB |  |
| **Max length** | |  | 2 |  |
| **Median length** | |  | 2 |  |
| **Mean length** | |  | 1.7928994 |  |
| **Min length**  Characters and Unicode | |  | 1 |  |
| **Total characters** | | 909 |  |  |
| **Distinct characters** | | 4 |  |  |
| **Distinct categories** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_character\_property#General\_Category)](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) | | [**?**](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) |
| **Distinct scripts** | | 1 [(https://en.wikipedia.org/wiki/Script\_(Unicode)#List\_of\_scripts\_in\_Unicode)](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) | | [**?**](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) |
| **Distinct blocks** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_block)](https://en.wikipedia.org/wiki/Unicode_block) | | [**?**](https://en.wikipedia.org/wiki/Unicode_block) |
| Unique | The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables. | | |  |
| **Unique** | 0 | | | **?** |
| **Unique (%)**  Sample | 0.0% | | |  |
| **1st row** | 15 | | |  |
| **2nd row** | 15 | | |  |
| **3rd row** | 15 | | |  |
| **4th row** | 5 | | |  |

**5th row** 15

Common Values

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 15 | 156 | 30.8% |
| 10 | 147 | 29.0% |
| 25 | 99 | 19.5% |
| 0 | 57 | 11.2% |

5 48 9.5%

|  |  |
| --- | --- |
|  |  |
|  |

Length

Histogram of lengths of the category

Common Values (Plot)

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 15 | 156 | 30.8% |
| 10 | 147 | 29.0% |
| 25 | 99 | 19.5% |
| 0 | 57 | 11.2% |

5 48 9.5%

Most occurring characters

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 1 | 303 | 33.3% |
| 5 | 303 | 33.3% |
| 0 | 204 | 22.4% |

2 99 10.9%

Most occurring categories

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 909 | 100.0% |
| Most frequent character per category  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| 1 | 303 | 33.3% |
| 5 | 303 | 33.3% |
| 0 | 204 | 22.4% |

2 99 10.9%

Most occurring scripts

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 909 | 100.0% |
| Most frequent character per script  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| 1 | 303 | 33.3% |
| 5 | 303 | 33.3% |
| 0 | 204 | 22.4% |

2 99 10.9%

Most occurring blocks

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 909 | 100.0% |
| Most frequent character per block  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| 1 | 303 | 33.3% |
| 5 | 303 | 33.3% |
| 0 | 204 | 22.4% |

2 99 10.9%

order\_type

Categorical

## HIGH CORRELATION (This variable has a high overall correlation with 1 fields: order\_payment) MISSING

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Distinct** | |  | 2 |  |
| **Distinct (%)** | |  | 0.4% |  |
| **Missing** | |  | 23 |  |
| **Missing (%)** | |  | 4.5% |  |
| **Memory size**  Length | |  | 4.1 KiB |  |
| **Max length** | |  | 7 |  |
| **Median length** | |  | 7 |  |
| **Mean length** | |  | 5.4628099 |  |
| **Min length**  Characters and Unicode | |  | 3 |  |
| **Total characters** | | 2644 |  |  |
| **Distinct characters** | | 10 |  |  |
| **Distinct categories** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_character\_property#General\_Category)](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) | | [**?**](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) |
| **Distinct scripts** | | 1 [(https://en.wikipedia.org/wiki/Script\_(Unicode)#List\_of\_scripts\_in\_Unicode)](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) | | [**?**](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) |
| **Distinct blocks** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_block)](https://en.wikipedia.org/wiki/Unicode_block) | | [**?**](https://en.wikipedia.org/wiki/Unicode_block) |
| Unique | The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables. | | |  |
| **Unique** | 0 | | | **?** |
| **Unique (%)**  Sample | 0.0% | | |  |
| **1st row** | COD | | |  |
| **2nd row** | COD | | |  |
| **3rd row** | COD | | |  |
| **4th row** | COD | | |  |

**5th row** COD

Common Values

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| Prepaid | 298 | 58.8% |
| COD | 186 | 36.7% |

(Missing) 23 4.5%

|  |  |
| --- | --- |
|  |  |
|  |

Length

Histogram of lengths of the category

Common Values (Plot)

|  |  |
| --- | --- |
|  |  |

**Value Count Frequency (%)**

61.6

%

prepaid

298

cod 186 38.4%

Most occurring characters

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| P | 298 | 11.3% |
| r | 298 | 11.3% |
| e | 298 | 11.3% |
| p | 298 | 11.3% |
| a | 298 | 11.3% |
| i | 298 | 11.3% |
| d | 298 | 11.3% |
| C | 186 | 7.0% |
| O | 186 | 7.0% |

D 186 7.0%

Most occurring categories

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 2644 | 100.0% |
| Most frequent character per category  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| P | 298 | 11.3% |
| r | 298 | 11.3% |
| e | 298 | 11.3% |
| p | 298 | 11.3% |
| a | 298 | 11.3% |
| i | 298 | 11.3% |
| d | 298 | 11.3% |
| C | 186 | 7.0% |
| O | 186 | 7.0% |

D 186 7.0%

Most occurring scripts

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 2644 | 100.0% |
| Most frequent character per script  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| P | 298 | 11.3% |
| r | 298 | 11.3% |
| e | 298 | 11.3% |
| p | 298 | 11.3% |
| a | 298 | 11.3% |
| i | 298 | 11.3% |
| d | 298 | 11.3% |
| C | 186 | 7.0% |
| O | 186 | 7.0% |

D 186 7.0%

Most occurring blocks

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 2644 | 100.0% |
| Most frequent character per block  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| P | 298 | 11.3% |
| r | 298 | 11.3% |
| e | 298 | 11.3% |
| p | 298 | 11.3% |
| a | 298 | 11.3% |
| i | 298 | 11.3% |
| d | 298 | 11.3% |
| C | 186 | 7.0% |
| O | 186 | 7.0% |

D 186 7.0%

order\_payment

Categorical

## HIGH CORRELATION (This variable has a high overall correlation with 1 fields: order\_type) MISSING

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Distinct** | |  | 5 |  |
| **Distinct (%)** | |  | 1.0% |  |
| **Missing** | |  | 17 |  |
| **Missing (%)** | |  | 3.4% |  |
| **Memory size**  Length | |  | 4.1 KiB |  |
| **Max length** | |  | 14 |  |
| **Median length** | |  | 11 |  |
| **Mean length** | |  | 8.6693878 |  |
| **Min length**  Characters and Unicode | |  | 3 |  |
| **Total characters** | | 4248 |  |  |
| **Distinct characters** | | 18 |  |  |
| **Distinct categories** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_character\_property#General\_Category)](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) | | [**?**](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) |
| **Distinct scripts** | | 1 [(https://en.wikipedia.org/wiki/Script\_(Unicode)#List\_of\_scripts\_in\_Unicode)](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) | | [**?**](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) |
| **Distinct blocks** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_block)](https://en.wikipedia.org/wiki/Unicode_block) | | [**?**](https://en.wikipedia.org/wiki/Unicode_block) |
| Unique | The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables. | | |  |
| **Unique** | 0 | | | **?** |
| **Unique (%)**  Sample | 0.0% | | |  |
| **1st row** | Cash | | |  |
| **2nd row** | Cash | | |  |
| **3rd row** | Debit Card | | |  |
| **4th row** | Debit Card | | |  |

**5th row** Credit Card

Common Values

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| Digital Wallet | 125 | 24.7% |
| Cash | 122 | 24.1% |
| Debit Card | 87 | 17.2% |
| Credit Card | 84 | 16.6% |
| EMI | 72 | 14.2% |

(Missing) 17 3.4%

Length

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | |  |  | | --- | --- | |  |  | |  | | | |
|  |
|  |  |  |

Histogram of lengths of the category

Common Values (Plot)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| card | 171 | 21.8% |
| digital | 125 | 15.9% |
| wallet | 125 | 15.9% |
| cash | 122 | 15.5% |
| debit | 87 | 11.1% |
| credit | 84 | 10.7% |

emi 72 9.2%

Most occurring characters

**Value Count Frequency (%)**

|  |  |  |  |
| --- | --- | --- | --- |
| a | 543 | 12.8% | |
| i | 421 | 9.9% |  |
| t | 421 | 9.9% |  |
| C | 377 | 8.9% |  |
| l | 375 | 8.8% |  |
|  | 296 |  | 7.0% |
| e | 296 |  | 7.0% |
| d | 255 |  | 6.0% |
| r | 255 |  | 6.0% |
| D | 212 |  | 5.0% |

Other values (8) 797 18.8%

Most occurring categories

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4248 | 100.0% |
| Most frequent character per category  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| a | 543 | 12.8% |
| i | 421 | 9.9% |
| t | 421 | 9.9% |
| C | 377 | 8.9% |
| l | 375 | 8.8% |
|  | 296 | 7.0% |
| e | 296 | 7.0% |
| d | 255 | 6.0% |
| r | 255 | 6.0% |
| D | 212 | 5.0% |

Other values (8) 797 18.8%

Most occurring scripts

**Value**

|  |  |  |  |
| --- | --- | --- | --- |
| (unknown) | 4248 | 100.0% | |
| Most frequent character per script  *(unknown)*  **Value** | **Count** | **Frequency (%)** | |
| a | 543 | 12.8% | |
| i | 421 | 9.9% |  |
| t | 421 | 9.9% |  |
| C | 377 | 8.9% |  |
| l | 375 | 8.8% |  |
|  | 296 |  | 7.0% |
| e | 296 |  | 7.0% |
| d | 255 |  | 6.0% |
| r | 255 |  | 6.0% |
| D | 212 |  | 5.0% |

Other values (8) 797 18.8%

Most occurring blocks

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 4248 | 100.0% |
| Most frequent character per block  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| a | 543 | 12.8% |
| i | 421 | 9.9% |
| t | 421 | 9.9% |
| C | 377 | 8.9% |
| l | 375 | 8.8% |
|  | 296 | 7.0% |
| e | 296 | 7.0% |
| d | 255 | 6.0% |
| r | 255 | 6.0% |
| D | 212 | 5.0% |

Other values (8) 797 18.8%

is\_expedited\_delivery

Boolean

## HIGH CORRELATION (This variable has a high overall correlation with 2 fields: customer\_satisfaction, delivery\_charges)

|  |  |  |
| --- | --- | --- |
| **Distinct** | 2 |  |
| **Distinct (%)** | 0.4% |  |
| **Missing** | 0 |  |
| **Missing (%)** | 0.0% |  |
| **Memory size** | 635.0 B |  |
| **Value** | **Count** | **Frequency (%)** |
| False | 256 | 50.5% |

True 251 49.5%

|  |  |
| --- | --- |
|  |  |

distance\_to\_nearest\_warehouse

Real number (ℝ)

|  |  |
| --- | --- |
| **Distinct** | 486 |
| **Distinct (%)** | 95.9% |
| **Missing** | 0 |
| **Missing (%)** | 0.0% |
| **Infinite** | 0 |
| **Infinite (%)** | 0.0% |
| **Mean** | 2.3745961 |
| **Minimum** | 0.1078 |
| **Maximum** | 94.9734 |
| **Zeros** | 0 |
| **Zeros (%)** | 0.0% |
| **Negative** | 0 |
| **Negative (%)** | 0.0% |

|  |  |
| --- | --- |
| **Minimum** | 0.1078 |
| **5-th percentile** | 0.32872 |
| **Q1** | 0.7517 |
| **median** | 1.0328 |

**Memory size** 4.1 KiB

Quantile statistics

|  |  |
| --- | --- |
| **Q3** | 1.40865 |
| **95-th percentile** | 1.99828 |
| **Maximum** | 94.9734 |
| **Range** | 94.8656 |
| **Interquartile range (IQR)** Descriptive statistics | 0.65695 |
| **Standard deviation** | 9.6736329 |
| **Coefficient of variation (CV)** | 4.0738015 |
| **Kurtosis** | 59.489756 |
| **Mean** | 2.3745961 |
| **Median Absolute Deviation (MAD)** | 0.3247 |
| **Skewness** | 7.6570126 |
| **Sum** | 1203.9202 |
| **Variance** | 93.579174 |

**Monotonicity** Not monotonic

**Histogram with fixed size bins** (bins=50)

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 0.2502 | 2 | 0.4% |
| 0.9083 | 2 | 0.4% |
| 0.9098 | 2 | 0.4% |
| 0.4 | 2 | 0.4% |
| 1.3171 | 2 | 0.4% |
| 0.8949 | 2 | 0.4% |
| 0.7686 | 2 | 0.4% |
| 0.8596 | 2 | 0.4% |
| 1.0406 | 2 | 0.4% |
| 1.0176 | 2 | 0.4% |

Other values (476) 487 96.1%

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 0.1078 | 1 | 0.2% |
| 0.1227 | 1 | 0.2% |
| 0.1282 | 1 | 0.2% |
| 0.1459 | 1 | 0.2% |
| 0.154 | 1 | 0.2% |
| 0.1542 | 1 | 0.2% |
| 0.1581 | 1 | 0.2% |
| 0.1702 | 1 | 0.2% |
| 0.1753 | 1 | 0.2% |

0.1928 1 0.2%

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 94.9734 | 2 | 0.4% |
| 76.3419 | 1 | 0.2% |
| 73.8324 | 1 | 0.2% |
| 68.8892 | 1 | 0.2% |
| 66.8512 | 1 | 0.2% |
| 66.6483 | 1 | 0.2% |
| 61.2935 | 1 | 0.2% |
| 48.5984 | 1 | 0.2% |
| 20.8952 | 1 | 0.2% |

2.7391 1 0.2%

customer\_satisfaction

Boolean

## HIGH CORRELATION (This variable has a high overall correlation with 1 fields: is\_expedited\_delivery)

|  |  |  |
| --- | --- | --- |
| **Distinct** | 2 |  |
| **Distinct (%)** | 0.4% |  |
| **Missing** | 0 |  |
| **Missing (%)** | 0.0% |  |
| **Memory size** | 635.0 B |  |
| **Value** | **Count** | **Frequency (%)** |
| True | 263 | 51.9% |

False 244 48.1%

|  |  |
| --- | --- |
|  |  |

age

Real number (ℝ)

|  |  |
| --- | --- |
| **Distinct** | 11 |
| **Distinct (%)** | 2.2% |
| **Missing** | 0 |
| **Missing (%)** | 0.0% |
| **Infinite** | 0 |
| **Infinite (%)** | 0.0% |
| **Mean** | 27.29783 |
| **Minimum** | 22 |
| **Maximum** | 32 |
| **Zeros** | 0 |
| **Zeros (%)** | 0.0% |
| **Negative** | 0 |
| **Negative (%)** | 0.0% |

|  |  |
| --- | --- |
| **Minimum** | 22 |
| **5-th percentile** | 23 |
| **Q1** | 25 |
| **median** | 27 |

**Memory size** 4.1 KiB

Quantile statistics

|  |  |  |
| --- | --- | --- |
| **Q3** |  | 30 |
| **95-th percentile** |  | 32 |
| **Maximum** |  | 32 |
| **Range** |  | 10 |
| **Interquartile range (IQR)** Descriptive statistics |  | 5 |
| **Standard deviation** | 2.9029178 |  |
| **Coefficient of variation (CV)** | 0.10634244 |  |
| **Kurtosis** | -1.2098478 |  |
| **Mean** | 27.29783 |  |
| **Median Absolute Deviation (MAD)** | 3 |  |
| **Skewness** | -0.040413102 |  |
| **Sum** | 13840 |  |
| **Variance** | 8.426932 |  |

**Monotonicity** Not monotonic

**Histogram with fixed size bins** (bins=11)

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 30 | 58 | 11.4% |
| 23 | 53 | 10.5% |
| 31 | 53 | 10.5% |
| 29 | 53 | 10.5% |
| 24 | 52 | 10.3% |
| 27 | 51 | 10.1% |
| 26 | 48 | 9.5% |
| 25 | 48 | 9.5% |
| 28 | 46 | 9.1% |

32 35 6.9%

**Value Count Frequency (%)**

|  |  |
| --- | --- |
| 22 10 | 2.0% |
| 23 53 | 10.5% |
| 24 52 | 10.3% |
| 25 48 | 9.5% |
| 26 48 | 9.5% |
| 27 51 | 10.1% |
| 28 46 | 9.1% |
| 29 53 | 10.5% |
| 30 58 | 11.4% |

31 53 10.5%

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 32 | 35 | 6.9% |
| 31 | 53 | 10.5% |
| 30 | 58 | 11.4% |
| 29 | 53 | 10.5% |
| 28 | 46 | 9.1% |
| 27 | 51 | 10.1% |
| 26 | 48 | 9.5% |
| 25 | 48 | 9.5% |
| 24 | 52 | 10.3% |

23 53 10.5%

warehouse\_area\_type

Categorical

## HIGH CORRELATION (This variable has a high overall correlation with 2 fields: warehouse\_name, warehouse\_size)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Distinct** | |  | 3 |  |
| **Distinct (%)** | |  | 0.6% |  |
| **Missing** | |  | 0 |  |
| **Missing (%)** | |  | 0.0% |  |
| **Memory size**  Length | |  | 4.1 KiB |  |
| **Max length** | |  | 11 |  |
| **Median length** | |  | 10 |  |
| **Mean length** | |  | 10.236686 |  |
| **Min length**  Characters and Unicode | |  | 10 |  |
| **Total characters** | | 5190 |  |  |
| **Distinct characters** | | 16 |  |  |
| **Distinct categories** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_character\_property#General\_Category)](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) | | [**?**](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) |
| **Distinct scripts** | | 1 [(https://en.wikipedia.org/wiki/Script\_(Unicode)#List\_of\_scripts\_in\_Unicode)](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) | | [**?**](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) |
| **Distinct blocks** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_block)](https://en.wikipedia.org/wiki/Unicode_block) | | [**?**](https://en.wikipedia.org/wiki/Unicode_block) |
| Unique | The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables. | | |  |
| **Unique** | 0 | | | **?** |
| **Unique (%)**  Sample | 0.0% | | |  |
| **1st row** | Commercial | | |  |
| **2nd row** | Industrial | | |  |
| **3rd row** | Commercial | | |  |
| **4th row** | Commercial | | |  |

**5th row** Commercial

Common Values

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| Industrial | 200 | 39.4% |
| Commercial | 187 | 36.9% |

Residential 120 23.7%

|  |  |
| --- | --- |
|  |  |
|  |

Length

Histogram of lengths of the category

Common Values (Plot)

|  |  |  |
| --- | --- | --- |
|  |  |  |

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| industrial | 200 | 39.4% |
| commercial | 187 | 36.9% |

residential 120 23.7%

Most occurring characters

**Value Count Frequency (%)**

|  |  |  |  |
| --- | --- | --- | --- |
| i | 627 | 12.1% | |
| a | 507 | 9.8% |  |
| l | 507 | 9.8% |  |
| e | 427 |  | 8.2% |
| r | 387 |  | 7.5% |
| m | 374 |  | 7.2% |
| n | 320 |  | 6.2% |
| d | 320 |  | 6.2% |
| s | 320 |  | 6.2% |
| t | 320 |  | 6.2% |

Other values (6) 1081 20.8%

Most occurring categories

**Value**

|  |  |  |  |
| --- | --- | --- | --- |
| (unknown) | 5190 | 100.0% | |
| Most frequent character per category  *(unknown)*  **Value** | **Count** | **Frequency (%)** | |
| i | 627 | 12.1% | |
| a | 507 | 9.8% |  |
| l | 507 | 9.8% |  |
| e | 427 |  | 8.2% |
| r | 387 |  | 7.5% |
| m | 374 |  | 7.2% |
| n | 320 |  | 6.2% |
| d | 320 |  | 6.2% |
| s | 320 |  | 6.2% |
| t | 320 |  | 6.2% |

Other values (6) 1081 20.8%

Most occurring scripts

**Value**

|  |  |  |  |
| --- | --- | --- | --- |
| (unknown) | 5190 | 100.0% | |
| Most frequent character per script  *(unknown)*  **Value** | **Count** | **Frequency (%)** | |
| i | 627 | 12.1% | |
| a | 507 | 9.8% |  |
| l | 507 | 9.8% |  |
| e | 427 |  | 8.2% |
| r | 387 |  | 7.5% |
| m | 374 |  | 7.2% |
| n | 320 |  | 6.2% |
| d | 320 |  | 6.2% |
| s | 320 |  | 6.2% |
| t | 320 |  | 6.2% |

Other values (6) 1081 20.8%

Most occurring blocks

**Value**

|  |  |  |  |
| --- | --- | --- | --- |
| (unknown) | 5190 | 100.0% | |
| Most frequent character per block  *(unknown)*  **Value** | **Count** | **Frequency (%)** | |
| i | 627 | 12.1% | |
| a | 507 | 9.8% |  |
| l | 507 | 9.8% |  |
| e | 427 |  | 8.2% |
| r | 387 |  | 7.5% |
| m | 374 |  | 7.2% |
| n | 320 |  | 6.2% |
| d | 320 |  | 6.2% |
| s | 320 |  | 6.2% |
| t | 320 |  | 6.2% |

Other values (6) 1081 20.8%

warehouse\_size

Categorical

## HIGH CORRELATION (This variable has a high overall correlation with 2 fields: warehouse\_area\_type, warehouse\_name)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Distinct** | |  | 3 |  |
| **Distinct (%)** | |  | 0.6% |  |
| **Missing** | |  | 0 |  |
| **Missing (%)** | |  | 0.0% |  |
| **Memory size**  Length | |  | 4.1 KiB |  |
| **Max length** | |  | 6 |  |
| **Median length** | |  | 5 |  |
| **Mean length** | |  | 5.3688363 |  |
| **Min length**  Characters and Unicode | |  | 5 |  |
| **Total characters** | | 2722 |  |  |
| **Distinct characters** | | 12 |  |  |
| **Distinct categories** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_character\_property#General\_Category)](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) | | [**?**](https://en.wikipedia.org/wiki/Unicode_character_property#General_Category) |
| **Distinct scripts** | | 1 [(https://en.wikipedia.org/wiki/Script\_(Unicode)#List\_of\_scripts\_in\_Unicode)](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) | | [**?**](https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode) |
| **Distinct blocks** | | 1 [(https://en.wikipedia.org/wiki/Unicode\_block)](https://en.wikipedia.org/wiki/Unicode_block) | | [**?**](https://en.wikipedia.org/wiki/Unicode_block) |
| Unique | The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables. | | |  |
| **Unique** | 0 | | | **?** |
| **Unique (%)**  Sample | 0.0% | | |  |
| **1st row** | Medium | | |  |
| **2nd row** | Large | | |  |
| **3rd row** | Medium | | |  |
| **4th row** | Medium | | |  |

**5th row** Medium

Common Values

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| Large | 200 | 39.4% |
| Medium | 187 | 36.9% |

Small 120 23.7%

|  |  |
| --- | --- |
|  |  |
|  |

Length

Histogram of lengths of the category

Common Values (Plot)

|  |  |  |
| --- | --- | --- |
|  |  |  |

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| large | 200 | 39.4% |
| medium | 187 | 36.9% |

small 120 23.7%

Most occurring characters

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| e | 387 | 14.2% |
| a | 320 | 11.8% |
| m | 307 | 11.3% |
| l | 240 | 8.8% |
| L | 200 | 7.3% |
| r | 200 | 7.3% |
| g | 200 | 7.3% |
| M | 187 | 6.9% |
| d | 187 | 6.9% |
| i | 187 | 6.9% |

Other values (2) 307 11.3%

Most occurring categories

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 2722 | 100.0% |
| Most frequent character per category  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| e | 387 | 14.2% |
| a | 320 | 11.8% |
| m | 307 | 11.3% |
| l | 240 | 8.8% |
| L | 200 | 7.3% |
| r | 200 | 7.3% |
| g | 200 | 7.3% |
| M | 187 | 6.9% |
| d | 187 | 6.9% |
| i | 187 | 6.9% |

Other values (2) 307 11.3%

Most occurring scripts

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 2722 | 100.0% |
| Most frequent character per script  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| e | 387 | 14.2% |
| a | 320 | 11.8% |
| m | 307 | 11.3% |
| l | 240 | 8.8% |
| L | 200 | 7.3% |
| r | 200 | 7.3% |
| g | 200 | 7.3% |
| M | 187 | 6.9% |
| d | 187 | 6.9% |
| i | 187 | 6.9% |

Other values (2) 307 11.3%

Most occurring blocks

**Value**

|  |  |  |
| --- | --- | --- |
| (unknown) | 2722 | 100.0% |
| Most frequent character per block  *(unknown)*  **Value** | **Count** | **Frequency (%)** |
| e | 387 | 14.2% |
| a | 320 | 11.8% |
| m | 307 | 11.3% |
| l | 240 | 8.8% |
| L | 200 | 7.3% |
| r | 200 | 7.3% |
| g | 200 | 7.3% |
| M | 187 | 6.9% |
| d | 187 | 6.9% |
| i | 187 | 6.9% |

Other values (2) 307 11.3%

total\_bill

Real number (ℝ)

## HIGH CORRELATION (This variable has a high overall correlation with 1 fields: order\_price)

|  |  |
| --- | --- |
| **Distinct** | 500 |
| **Distinct (%)** | 98.6% |
| **Missing** | 0 |
| **Missing (%)** | 0.0% |
| **Infinite** | 0 |
| **Infinite (%)** | 0.0% |
| **Mean** | 21843.925 |
| **Minimum** | 68.24 |
| **Maximum** | 853012.72 |
| **Zeros** | 0 |
| **Zeros (%)** | 0.0% |
| **Negative** | 0 |
| **Negative (%)** | 0.0% |

|  |  |
| --- | --- |
| **Minimum** | 68.24 |
| **5-th percentile** | 2240.996 |
| **Q1** | 6000.675 |

**Memory size** 4.1 KiB

Quantile statistics

|  |  |
| --- | --- |
| **median** | 11059.89 |
| **Q3** | 17662.385 |
| **95-th percentile** | 28107.326 |
| **Maximum** | 853012.72 |
| **Range** | 852944.48 |
| **Interquartile range (IQR)** Descriptive statistics | 11661.71 |
| **Standard deviation** | 74053.432 |
| **Coefficient of variation (CV)** | 3.3901156 |
| **Kurtosis** | 70.540112 |
| **Mean** | 21843.925 |
| **Median Absolute Deviation (MAD)** | 5620.28 |
| **Skewness** | 8.1408405 |
| **Sum** | 11074870 |
| **Variance** | 5.4839108 × 109 |

**Monotonicity** Not monotonic

**Histogram with fixed size bins** (bins=50)

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 12813.21 | 2 | 0.4% |
| 5615.39 | 2 | 0.4% |
| 5931.25 | 2 | 0.4% |
| 6508.63 | 2 | 0.4% |
| 2829.97 | 2 | 0.4% |
| 765.58 | 2 | 0.4% |
| 13683.9 | 2 | 0.4% |
| 7231.45 | 1 | 0.2% |
| 11975.07 | 1 | 0.2% |
| 3785.32 | 1 | 0.2% |

Other values (490) 490 96.6%

**Value Count Frequency (%)**

|  |  |  |
| --- | --- | --- |
| 68.24 | 1 | 0.2% |
| 74.84 | 1 | 0.2% |
| 75.93 | 1 | 0.2% |
| 78.28 | 1 | 0.2% |
| 101.29 | 1 | 0.2% |
| 530.54 | 1 | 0.2% |
| 765.58 | 2 | 0.4% |
| 915 | 1 | 0.2% |
| 961.91 | 1 | 0.2% |

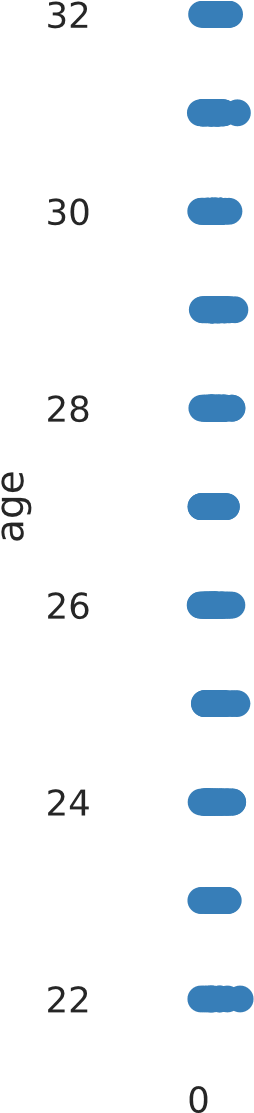
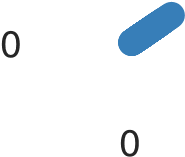
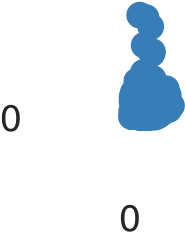
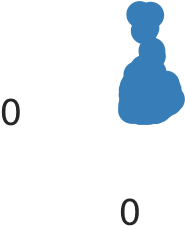
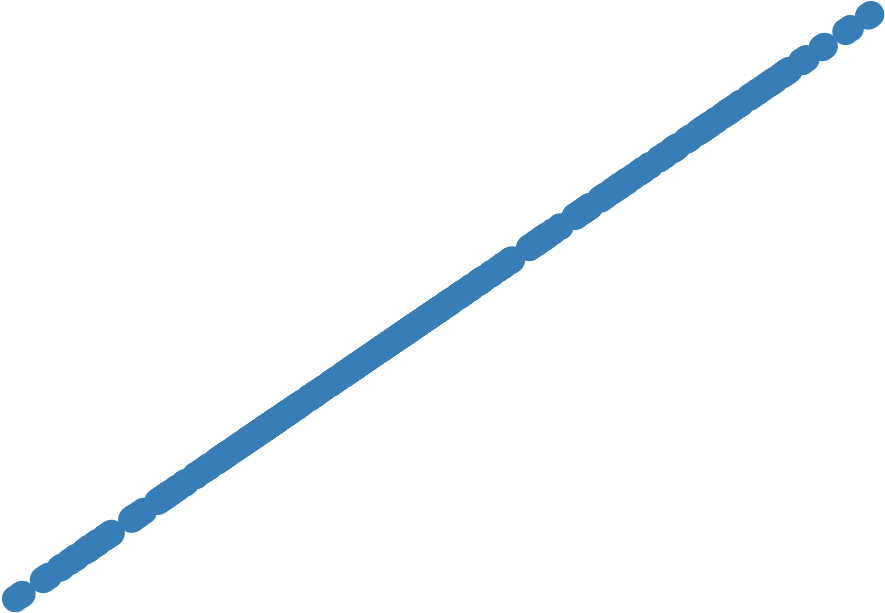
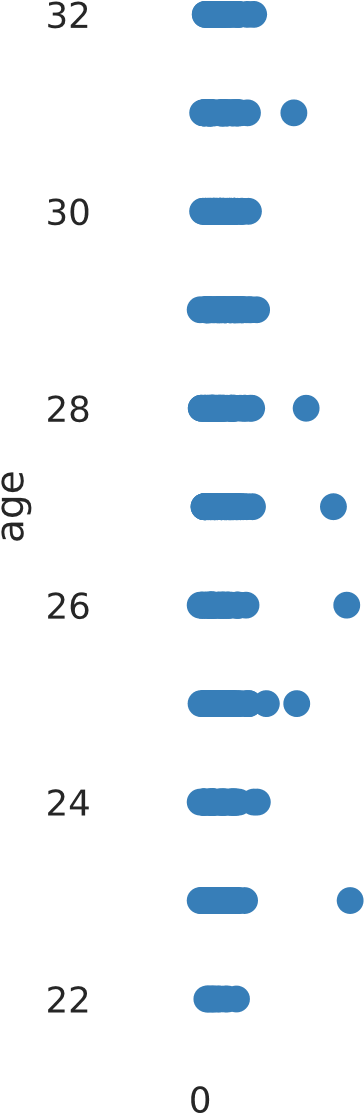
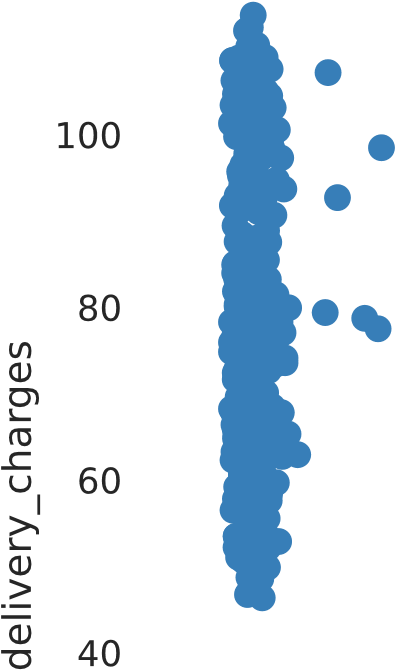
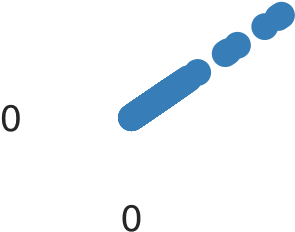
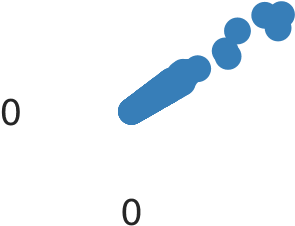
971.3 1 0.2%

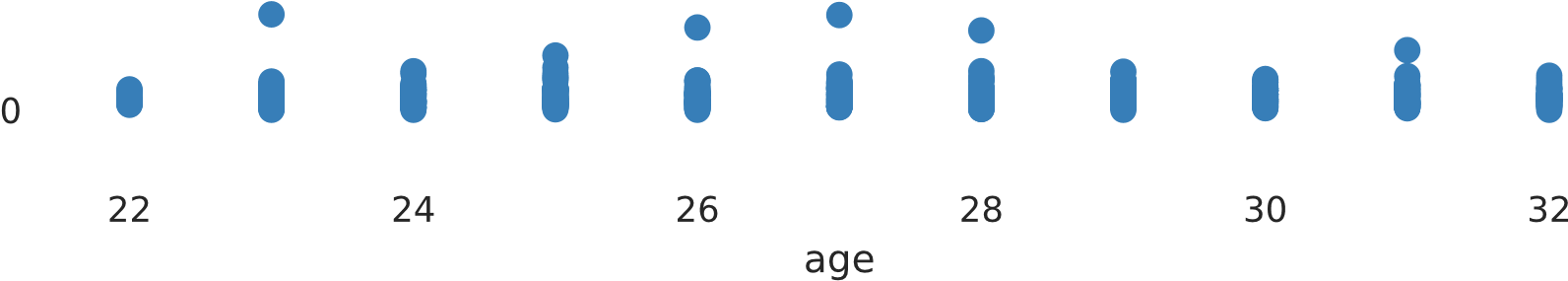
**Value Count Frequency (%)**

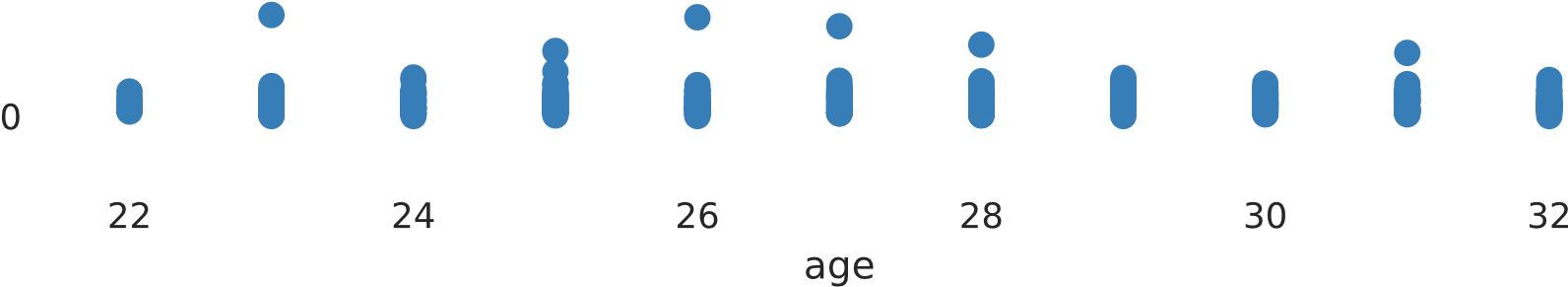
|  |  |  |
| --- | --- | --- |
| 853012.72 | 1 | 0.2% |
| 736121.5 | 1 | 0.2% |
| 589729.3 | 1 | 0.2% |
| 567457.91 | 1 | 0.2% |
| 555931.33 | 1 | 0.2% |
| 504686.82 | 1 | 0.2% |
| 386825.88 | 1 | 0.2% |
| 348158.1 | 1 | 0.2% |
| 346389.31 | 1 | 0.2% |

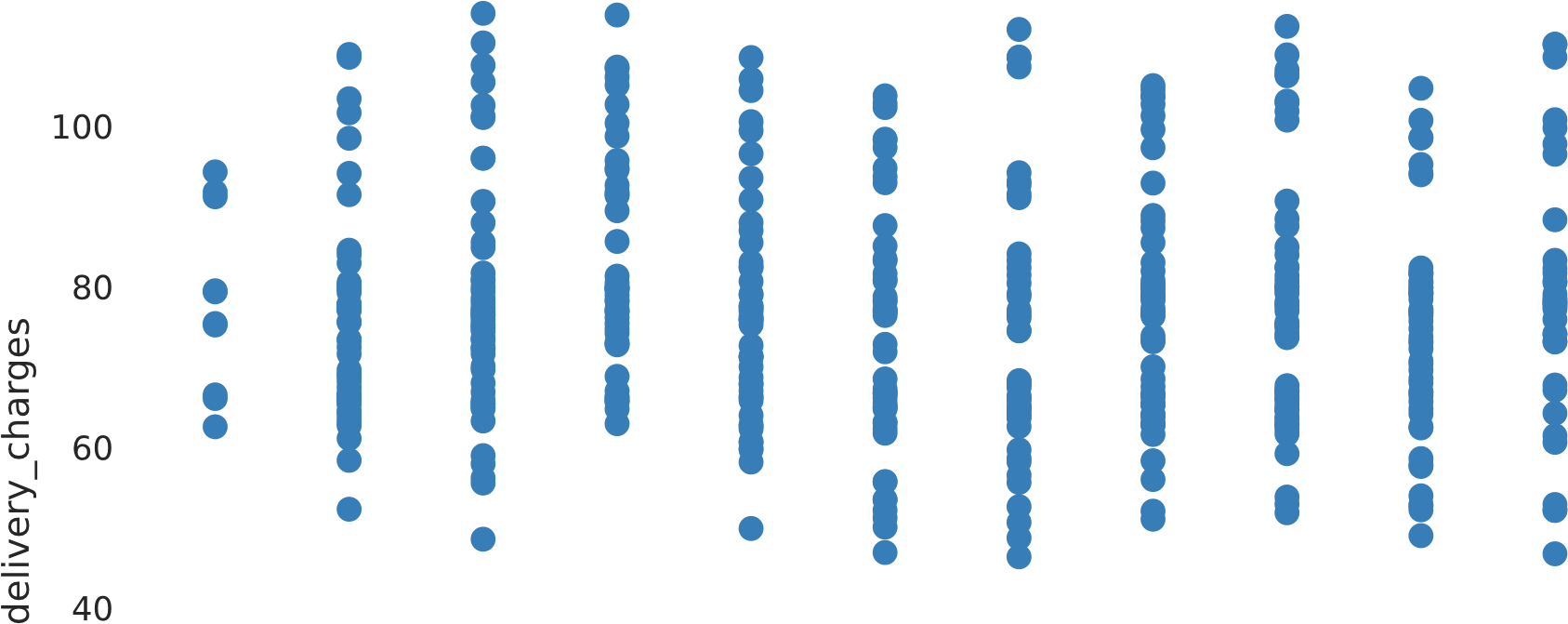
83920.38 1 0.2%

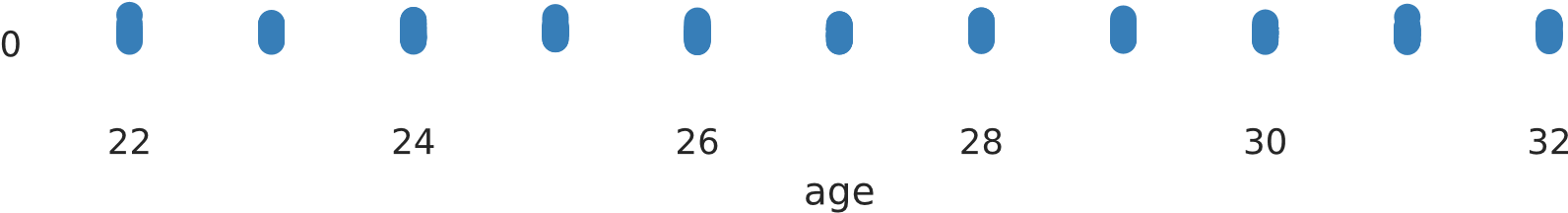
# Interactions

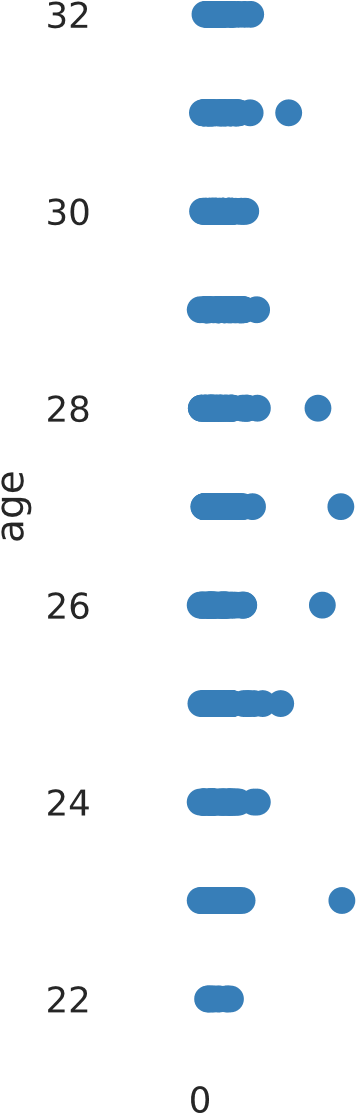
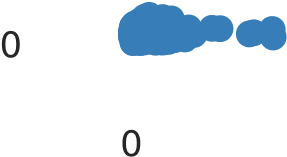
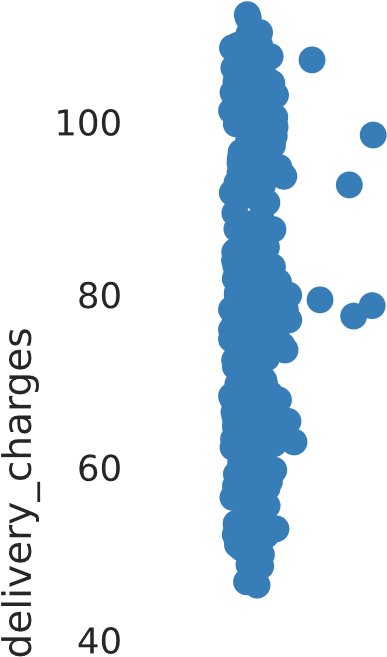
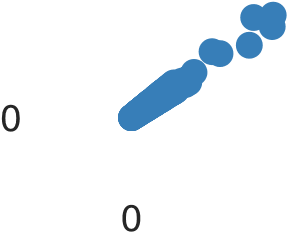
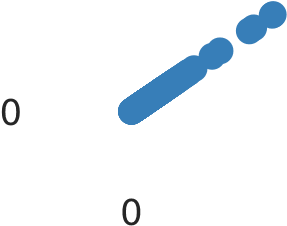












# Correlations



**age coupon\_discount customer\_satisfaction delivery\_charges distance\_to\_nearest\_warehouse ed**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **age** | 1.000 | 0.000 | 0.000 | -0.034 | -0.049 |
| **coupon\_discount** | 0.000 | 1.000 | 0.000 | -0.169 | 0.040 |
| **customer\_satisfaction** | 0.000 | 0.000 | 1.000 | 0.294 | 0.051 |
| **delivery\_charges** | -0.034 | -0.169 | 0.294 | 1.000 | 0.225 |
| **distance\_to\_nearest\_warehouse** | -0.049 | 0.040 | 0.051 | 0.225 | 1.000 |
| **education** | 0.052 | 0.055 | 0.011 | 0.000 | 0.000 |
| **is\_expedited\_delivery** | 0.000 | 0.183 | 0.601 | 0.608 | 0.044 |
| **marital\_status** | 0.000 | 0.000 | 0.000 | 0.108 | 0.052 |
| **occupation** | 0.000 | 0.014 | 0.068 | 0.000 | 0.153 |

0.0

0.0

0.0

0.0

0.0

1.0

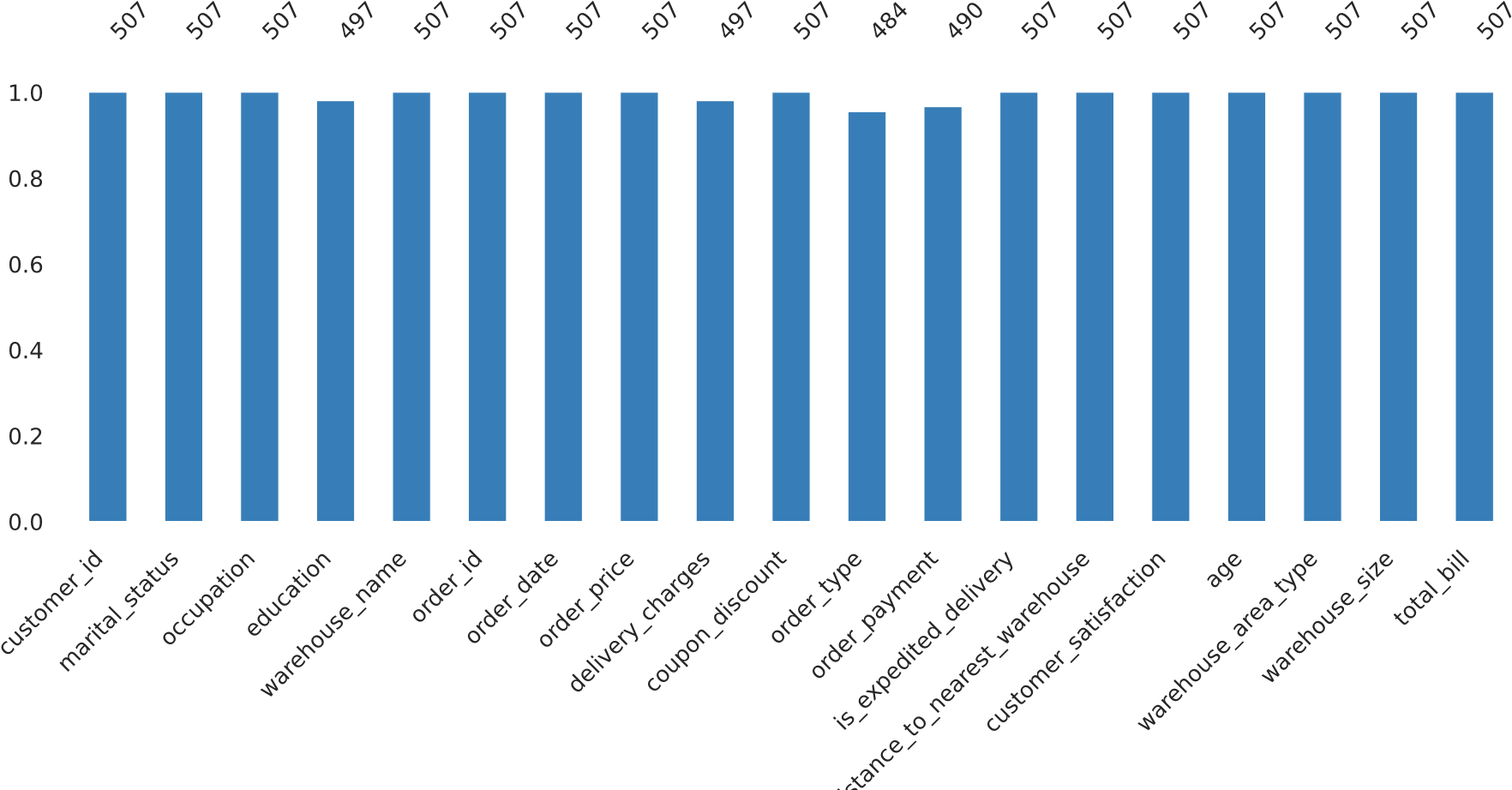
0.0

0.0

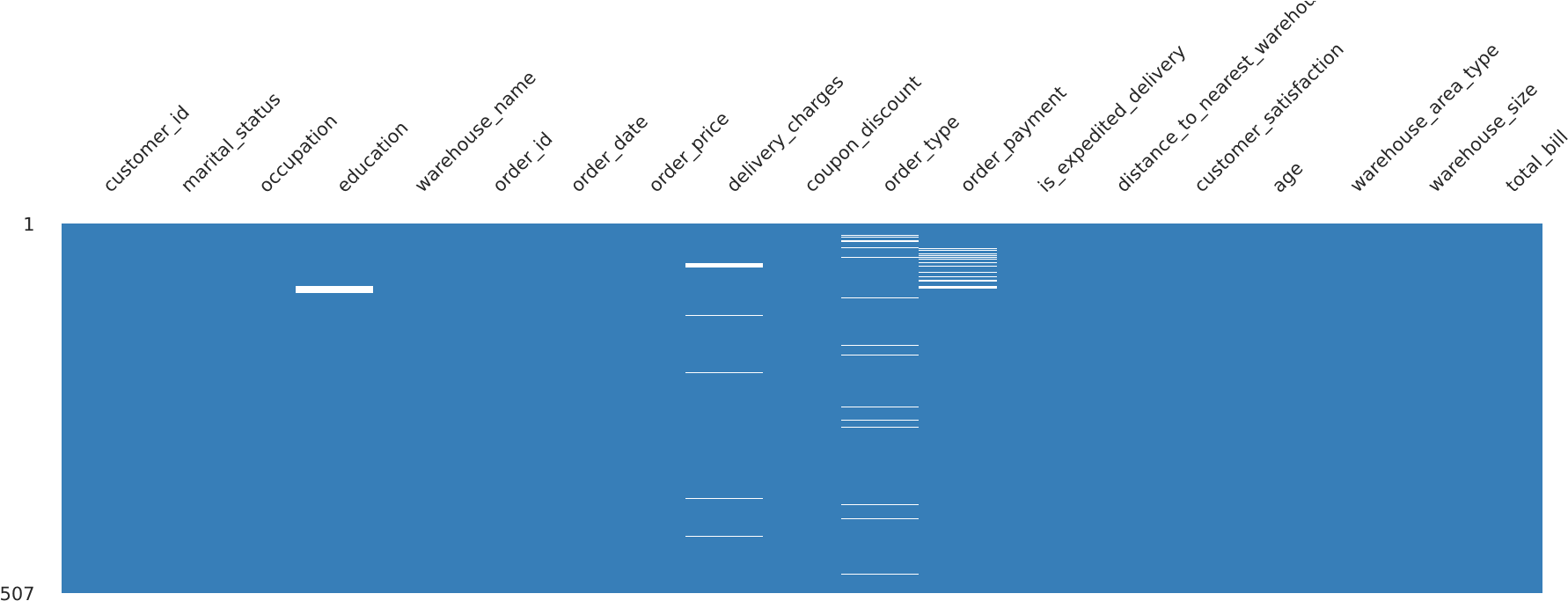
0.5

**order\_payment** 0.000 0.160 0.346 0.000 0.067 0.1

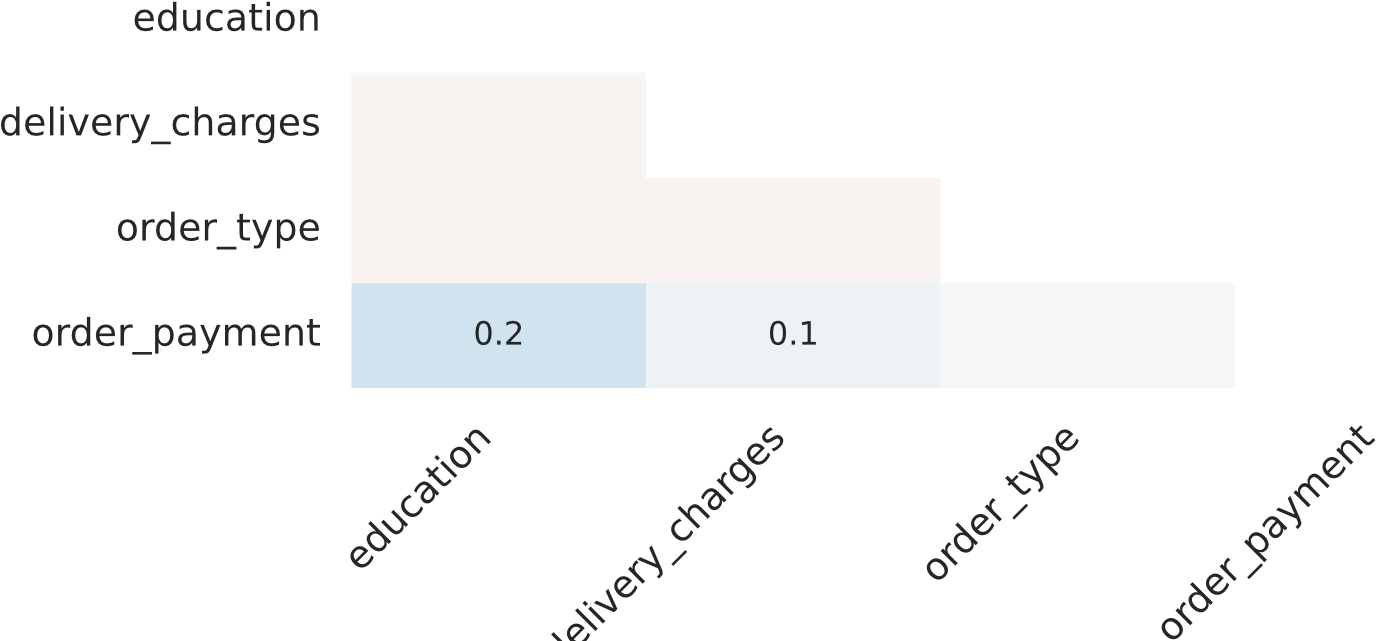
# Missing values



A simple visualization of nullity by column.



Nullity matrix is a data-dense display which lets you quickly visually pick out patterns in data completion.



The correlation heatmap measures nullity correlation: how strongly the presence or absence of one variable affects the presence of another.

# Sample

**customer\_id marital\_status occupation education warehouse\_name order\_id order\_date order\_price delivery\_charges cou**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** ID0571731474 | Yes | Manager | Masters | Nickolson | ORD001041 | 13-06-2021 | 8725 | 66.86 | 15 |
| **1** ID6167441063 | Yes | Clerical | School | Thompson | ORD004361 | 02-06-2021 | 22220 | 53.80 | 15 |
| **2** ID6221450539 | Yes | Lawyer | Masters | Nickolson | ORD004602 | 17-06-2021 | 10460 | 65.47 | 15 |
| **3** ID2383211221 | Yes | Manager | High School | Nickolson | ORD004848 | 13-06-2021 | 15070 | 79.13 | 5 |
| **4** ID0248129431 | No | Student | School | Nickolson | ORD005232 | 09-07-2021 | 5905 | 95.22 | 15 |
| **5** ID2222040558 | Yes | Clerical | High School | Thompson | ORD005286 | 21-07-2021 | 19560 | 82.44 | 10 |
| **6** ID5002592919 | Yes | Doctor | PhD | Nickolson | ORD007249 | 21-05-2021 | 6535 | 60.64 | 15 |
| **7** ID0247018750 | Yes | Blue Collar | High School | Thompson | ORD007904 | 29-05-2021 | 6865 | 96.00 | 15 |
| **8** ID3130999251 | No | Blue Collar | School | Thompson | ORD008127 | 30-07-2021 | 6770 | 77.13 | 5 |

**9** ID2187983888 Yes Professional Bachelors Nickolson ORD010057 18-05-2021 3085 53.47 10

**customer\_id marital\_status occupation education warehouse\_name order\_id order\_date order\_price delivery\_charges c**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **497** ID2096325801 | Yes | Clerical | High School | Thompson | ORD498207 | 10-06-2021 | 3750 | 80.04 |
| **498** ID0846545708 | No | Professional | Bachelors | Bakers | ORD498590 | 18-06-2021 | 6630 | 67.00 |
| **499** ID0052450483 | Yes | Lawyer | Masters | Nickolson | ORD499540 | 23-05-2021 | 14685 | 59.86 |
| **500** ID5002592919 | Yes | Doctor | PhD | Nickolson | ORD007249 | 21-05-2021 | 6535 | 60.64 |
| **501** ID0247018750 | Yes | Blue Collar | High School | Thompson | ORD007904 | 29-05-2021 | 6865 | 96.00 |
| **502** ID3130999251 | No | Blue Collar | School | Thompson | ORD008127 | 30-07-2021 | 6770 | 77.13 |
| **503** ID2187983888 | Yes | Professional | Bachelors | Nickolson | ORD010057 | 18-05-2021 | 3085 | 53.47 |
| **504** ID3516648756 | Yes | Blue Collar | High School | Nickolson | ORD010677 | 18-07-2021 | 730 | 108.58 |
| **505** ID0363973195 | Yes | Blue Collar | High School | Bakers | ORD010913 | 01-06-2021 | 14175 | 55.71 |

1

2

1

1

1

5

1

1

1

**506** ID6167417972 Yes Blue Collar High School Thompson ORD014742 29-05-2021 15130 66.90 1

# Duplicate rows

Most frequently occurring **customer\_id marital\_status occupation education warehouse\_name order\_id order\_date order\_price delivery\_charges coupon**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** ID0247018750 | Yes | Blue Collar | High School | Thompson | ORD007904 | 29-05-2021 | 6865 | 96.00 | 15 |
| **1** ID0363973195 | Yes | Blue Collar | High School | Bakers | ORD010913 | 01-06-2021 | 14175 | 55.71 | 10 |
| **2** ID2187983888 | Yes | Professional | Bachelors | Nickolson | ORD010057 | 18-05-2021 | 3085 | 53.47 | 10 |
| **3** ID3130999251 | No | Blue Collar | School | Thompson | ORD008127 | 30-07-2021 | 6770 | 77.13 | 5 |
| **4** ID3516648756 | Yes | Blue Collar | High School | Nickolson | ORD010677 | 18-07-2021 | 730 | 108.58 | 10 |
| **5** ID5002592919 | Yes | Doctor | PhD | Nickolson | ORD007249 | 21-05-2021 | 6535 | 60.64 | 15 |

**6** ID6167417972 Yes Blue Collar High School Thompson ORD014742 29-05-2021 15130 66.90 10

Report generated by [YData (https://ydata.ai/?utm\_source=opensource&utm\_medium=pandasprofiling&utm\_campaign=report).](https://ydata.ai/?utm_source=opensource&utm_medium=pandasprofiling&utm_campaign=report)