



24ti x 24ti x Pro x Rec x Di x Cod x (5) x 127 x kun x Hal x Hor x Unt x D-T x

Not secure | dell-05:40000/dtale/main/1

Gmail YouTube Maps News Translate

	enrollee_id	city	city_development_index	gender	relevent_experience
0	32403	city_41	0.83	Male	Has relevent experien
1	9858	city_103	0.92	Female	Has relevent experien
2	31806	city_21	0.62	Male	No relevent experien
3	27385	city_13	0.83	Male	Has relevent experien
4	27724	city_103	0.92	Male	Has relevent experien
5	217	city_23	0.90	Male	No relevent experien
6	21465	city_21	0.62	nan	Has relevent experien
7	27302	city_160	0.92	Female	Has relevent experien
8	12994	city_173	0.88	Male	Has relevent experien
9	16287	city_21	0.62	Male	Has relevent experien
10	10856	city_103	0.92	Male	Has relevent experien
11	9272	city_90	0.70	Male	Has relevent experien
12	14249	city_46	0.76	Male	Has relevent experien
13	24372	city_98	0.95	nan	Has relevent experien
14	14070	city_103	0.92	nan	No relevent experien
15	24914	city_21	0.62	nan	Has relevent experien
16	7865	city_21	0.62	Male	Has relevent experien
17	7463	city_13	0.83	Male	Has relevent experien
18	21514	city_21	0.62	nan	Has relevent experien
19	29033	city_21	0.62	Male	No relevent experien
20	15359	city_103	0.92	nan	No relevent experien
21	16001	city_103	0.92	nan	Has relevent experien
22	25202	city_21	0.62	Male	Has relevent experien

city_development_index (float64)

DescribeHistogramCategoriesQ-Q Plot

Total Rows: 2,129

- Count (non-nan): 1,621
- Count (missing): 508
- % Missing: 23.86

Most Frequent: Male

- Frequency: 1460

Unique: 3

Characters

- Min # Chars: 4
- Average # Chars: 4.18
- Max # Chars: 6
- STD # Chars: 0.57
- Rows w/ Spaces: 0
- Rows w/ Accent Chars: 0
- Rows w/ Numeric Chars: 0
- Rows w/ Uppercase Chars: 1621
- Rows w/ Lowercase Chars: 1621
- Rows w/ Punctuation: 0
- Rows Starting w/ Space: 0
- Rows Ending w/ Space: 0
- Rows w/ Multi Spacing: 0
- Rows w/ Hidden Chars: 0

Words

- Min # Words: 1
- Average # Words: 1
- Max # Words: 1

Windows taskbar with icons for File Explorer, Edge, Mail, Teams, VS Code, Word, PowerPoint, and various background applications.

System tray showing 30°C, Smoke, network status, ENG, and date/time: 11:53:37 16-10-2022.

24ti x 24ti x Pro x Rec x Di x Cod x (5) x 127 x kun x Hal x Hor x Uni x D-T x

Not secure | dell-05:40000/dtale/main/1

Update

Gmail YouTube Maps News Translate

13	enrollee_id	city	city_developm
2129			
0	32403	city_41	
1	9858	city_103	
2	31806	city_21	
3	27385	city_13	
4	27724	city_103	
5	217	city_23	
6	21465	city_21	
7	27302	city_160	
8	12994	city_173	
9	16287	city_21	
10	10856	city_103	
11	9272	city_90	
12	14249	city_46	
13	24372	city_98	
14	14070	city_103	
15	24914	city_21	
16	7865	city_21	
17	7463	city_13	
18	21514	city_21	
19	29033	city_21	
20	15359	city_103	
21	16001	city_103	
22	25202	city_21	

Pearson Correlation Matrix

(Click on any cell to view the details of that correlation)

View Correlation(s) For

city_developmen... x

 vs.

training_hours x

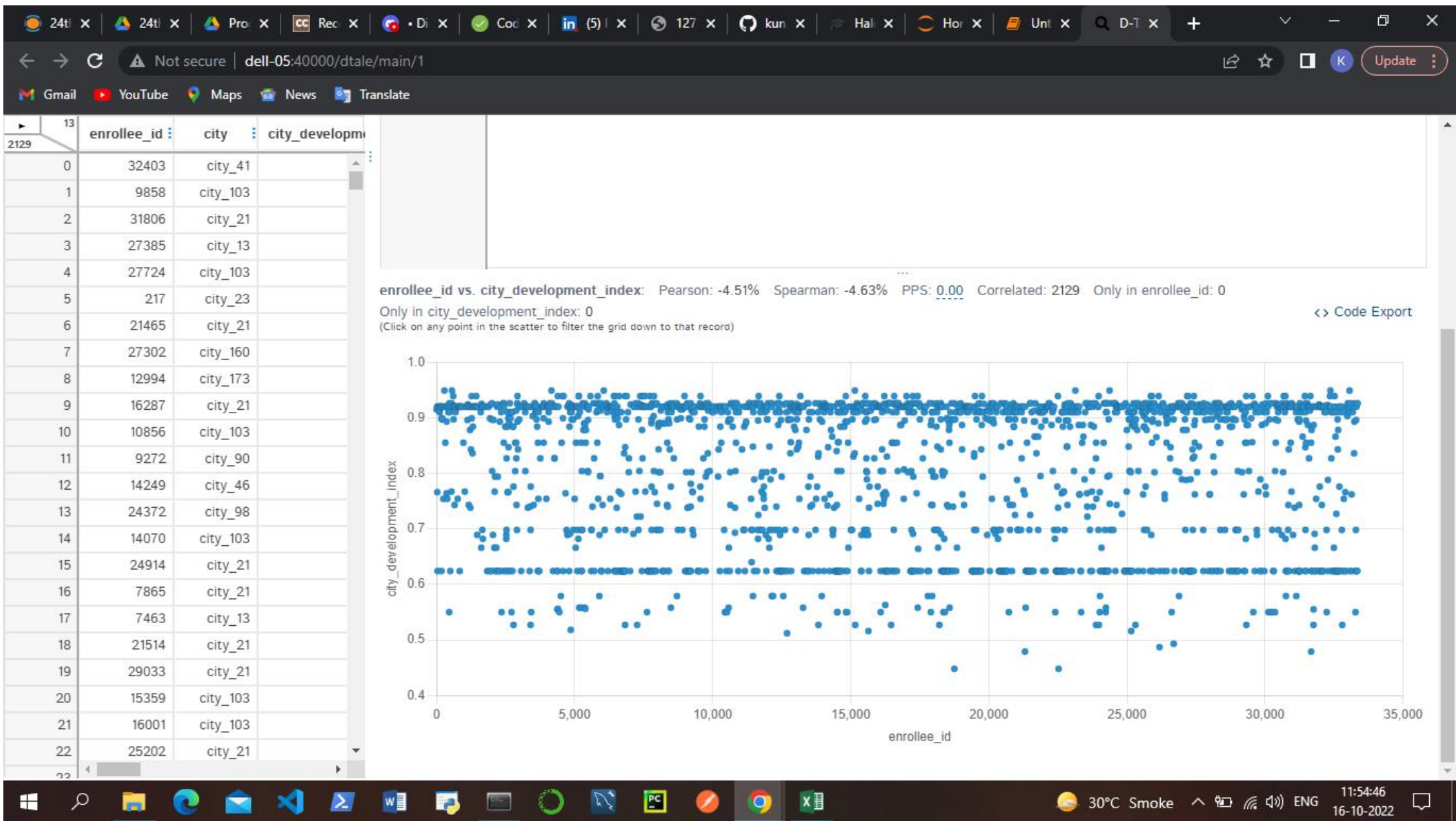
Encode Strings? ☐

	training_hours
city_developmen...	-0.02

enrollee_id vs. city_development_index: Pearson: -4.51% Spearman: -4.63% PPS: 0.00 Correlated: 2129 Only in enrollee_id: 0
Only in city_development_index: 0
(Click on any point in the scatter to filter the grid down to that record)

<> Code Export

Windows taskbar with icons for File Explorer, Edge, Mail, VS Code, PyCharm, Word, Excel, and system tray showing 30°C, Smoke, and date 16-10-2022.



24ti x 24ti x Pro x CC Rec x Di x Cod x in (5) x 127 x kun x Hal x Hor x Unt x D-T x +

Not secure | dell-05:40000/dtale/main/1

Gmail YouTube Maps News Translate

	enrollee_id	city	city_development_index	gender	relevent_experience	enrolled_university	education_level	major_discipline	experience	company_size	comp
0	32403	city_41	0.83	Male	Has relevent experience	Full time course	Graduate	STEM	9	<10	
1	9858	city_103	0.92	Female	Has relevent experience	no_enrollment	Graduate	STEM	5	nan	
2	31806	city_21	0.62	Male	No relevent experience	no_enrollment	High School	nan	<1	nan	
3	27385	city_13	0.83	Male	Has relevent experience	no_enrollment	Masters	STEM	11	10/49	
4	27724	city_103	0.92	Male	Has relevent experience	no_enrollment	Graduate	STEM	>20	10000+	
5	217	city_23	0.90	Male	No relevent experience	Part time course	Masters	STEM	10	nan	
6	21465	city_21	0.62	nan	Has relevent experience	no_enrollment	Graduate	STEM	<1	100-500	
7	27302	city_160	0.92	Female	Has relevent experience	no_enrollment	Graduate	STEM	>20	nan	
8	12994	city_173	0.88	Male	Has relevent experience	no_enrollment	Graduate	STEM	14	nan	
9	16287	city_21	0.62	Male	Has relevent experience	Full time course	Graduate	nan	3	50-99	
10	10856	city_103	0.92	Male	Has relevent experience	no_enrollment	Masters	Other	>20	nan	
11	9272	city_90	0.70	Male	Has relevent experience	no_enrollment	Graduate	STEM	20	10/49	
12	14249	city_46	0.76	Male	Has relevent experience	no_enrollment	Graduate	STEM	8	100-500	
13	24372	city_98	0.95	nan	Has relevent experience	no_enrollment	Masters	STEM	4	100-500	
14	14070	city_103	0.92	nan	No relevent experience	no_enrollment	Graduate	STEM	5	nan	
15	24914	city_21	0.62	nan	Has relevent experience	Full time course	Graduate	STEM	13	1000-4999	
16	7865	city_21	0.62	Male	Has relevent experience	no_enrollment	Masters	STEM	4	100-500	
17	7463	city_13	0.83	Male	Has relevent experience	no_enrollment	Masters	Business Degree	2	50-99	
18	21514	city_21	0.62	nan	Has relevent experience	no_enrollment	Graduate	STEM	6	nan	
19	29033	city_21	0.62	Male	No relevent experience	Full time course	nan	nan	2	nan	
20	15359	city_103	0.92	nan	No relevent experience	Full time course	Graduate	STEM	2	nan	
21	16001	city_103	0.92	nan	Has relevent experience	no_enrollment	Graduate	STEM	7	10000+	
22	25202	city_21	0.62	Male	Has relevent experience	no_enrollment	Graduate	STEM	6	1000-4999	

30°C Smoke 12:07:15 16-10-2022

24ti x 24ti x Pro x Rec x Di x Cod x (5) x 127 x kun x Hal x Hor x Unt x D-T x

Not secure | dell-05:40000/dtale/main/1

Gmail YouTube Maps News Translate

	enrollee_id	city	city_development_index	gender	relevent_experience
0	32403	city_41	0.83	Male	Has relevent experien
1	9858	city_103	0.92	Female	Has relevent experien
2	31806	city_21	0.62	Male	No relevent experien
3	27385	city_13	0.83	Male	Has relevent experien
4	27724	city_103	0.92	Male	Has relevent experien
5	217	city_23	0.90	Male	No relevent experien
6	21465	city_21	0.62	nan	Has relevent experien
7	27302	city_160	0.92	Female	Has relevent experien
8	12994	city_173	0.88	Male	Has relevent experien
9	16287	city_21	0.62	Male	Has relevent experien
10	10856	city_103	0.92	Male	Has relevent experien
11	9272	city_90	0.70	Male	Has relevent experien
12	14249	city_46	0.76	Male	Has relevent experien
13	24372	city_98	0.95	nan	Has relevent experien
14	14070	city_103	0.92	nan	No relevent experien
15	24914	city_21	0.62	nan	Has relevent experien
16	7865	city_21	0.62	Male	Has relevent experien
17	7463	city_13	0.83	Male	Has relevent experien
18	21514	city_21	0.62	nan	Has relevent experien
19	29033	city_21	0.62	Male	No relevent experien
20	15359	city_103	0.92	nan	No relevent experien
21	16001	city_103	0.92	nan	Has relevent experien
22	25202	city_21	0.62	Male	Has relevent experien

experience (string)

(Use ↑ ↓ buttons to switch columns)

DescribeWord Value CountsValue CountsCode Export

(Use ← → buttons to switch charts)

Value (Please edit)100Ordinal (Choose Col/Agg)Select...Sum

CleanerRemove Hidden Characters x

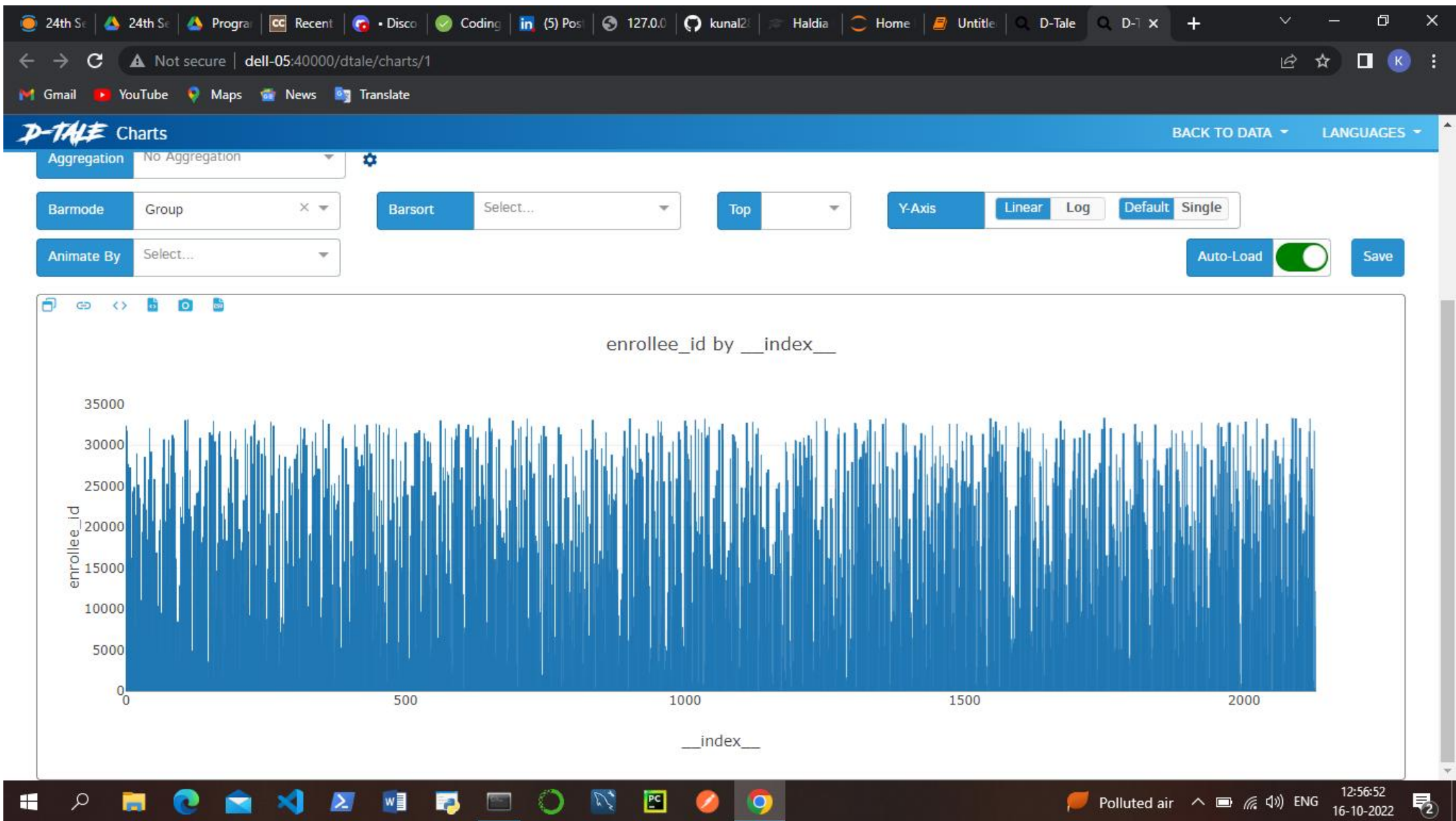
Frequency

Value

Unique Row Values (top 100 most common):>20 (383), 5 (163), 3 (154), 4 (145), 6 (130), 2 (128), 7 (116), 9 (113), 10 (96), 11 (86), 8 (82), <1 (74), 16 (68), 15 (68), 1 (68), 14 (68), 13 (68), 12 (68), 17 (68), 19 (68), 18 (68), 20 (68)

Windows Taskbar

30°C Smoke 12:07:59 16-10-2022



24th x 24th x Prog x CC Rece x Dis x Codi x (5) P x 127.0 x kua x Hald x Hom x Unti x +

localhost:8888/notebooks/Untitled26.ipynb

Gmail YouTube Maps News Translate

jupyter Untitled26 Last Checkpoint: 2 hours ago (autosaved)



Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted

Python 3 (ipykernel)

Run Cell Code

```
15 from pandas_profiling.config import Settings

~\anaconda3\lib\site-packages\statsmodels\graphics\tsaplots.py in <module>
9
10 from statsmodels.graphics import utils
--> 11 from statsmodels.tsa.stattools import acf, pacf
12
13

~\anaconda3\lib\site-packages\statsmodels\tsa\stattools.py in <module>
6 from statsmodels.compat.numpy import lstsq
7 from statsmodels.compat.pandas import deprecate_kwarg
----> 8 from statsmodels.compat.python import lzip, Literal
9 from statsmodels.compat.scipy import _next_regular
10

ImportError: cannot import name 'Literal' from 'statsmodels.compat.python' (C:\Users\A\anaconda3\lib\site-packages\statsmodels\compat\python.py)
```

In []:

Windows taskbar with icons for File Explorer, Edge, Mail, VS Code, JupyterLab, Word, PowerPoint, and Chrome. System tray shows 30°C, Smoke, and date/time: 13:17:48 16-10-2022.



Get updates, docs & report issues here
Created & maintained by [Francois Bertrand](#)
Graphic design by [Jean-Francois Hains](#)

city_development_index

NO COMPARISON TARGET

2129	ROWS
0	DUPLICATES
1.4 MB	RAM
13	FEATURES
9	CATEGORICAL
3	NUMERICAL
1	TEXT

ASSOCIATIONS

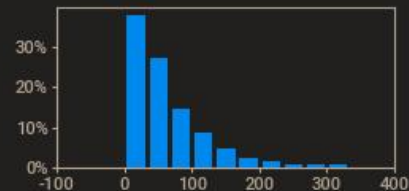
city_development_index
Avg. training_hours

training_hours

VALUES: 2,129 (100%)
MISSING: —
DISTINCT: 223 (10%)
ZEREOES: —

MAX 334
95% 192
Q3 86
AVG 65
MEDIAN 47
Q1 23
5% 7
MIN 1

RANGE 333
IQR 63.0
STD 60.2
VAR 3,629
KURT. 4.06
SKEW 1.88
SUM 138k

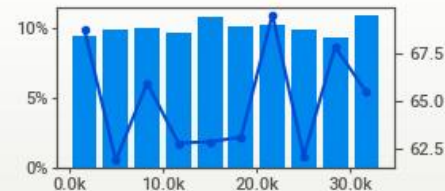


enrollee_id

VALUES: 2,129 (100%)
MISSING: —
DISTINCT: 2,129 (100%)
ZEREOES: —

MAX 33,353
95% 31,760
Q3 25,129
AVG 16,862
MEDIAN 16,816
Q1 8,562
5% 1,653
MIN 3

RANGE 33,350
IQR 16,567
STD 9,577
VAR 91.7M
KURT. -1.17
SKEW -0.015
SUM 35.9M



city

VALUES: 2,129 (100%)
MISSING: —
DISTINCT: 108 (5%)

473 22% city_103
318 15% city_21
168 8% city_16
155 7% city_114

2 city

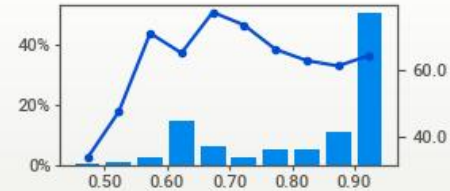
VALUES: 2,129 (100%)
MISSING: ---
DISTINCT: 108 (5%)

473	22%	city_103
318	15%	city_21
168	8%	city_16
155	7%	city_114
113	5%	city_160
60	3%	city_136
39	2%	city_67
803	38%	(Other)

3 city_development_index

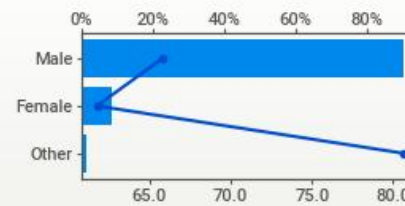
VALUES: 2,129 (100%)
MISSING: ---
DISTINCT: 81 (4%)
ZEROES: ---

MAX	0.949	RANGE	0.501
95%	0.926	IQR	0.222
Q3	0.920	STD	0.125
MEDIAN	0.903	VAR	0.016
AVG	0.825	KURT.	-0.686
Q1	0.698	SKEW	-0.923
5%	0.624	SUM	1,756
MIN	0.448		



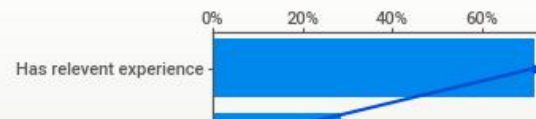
4 gender

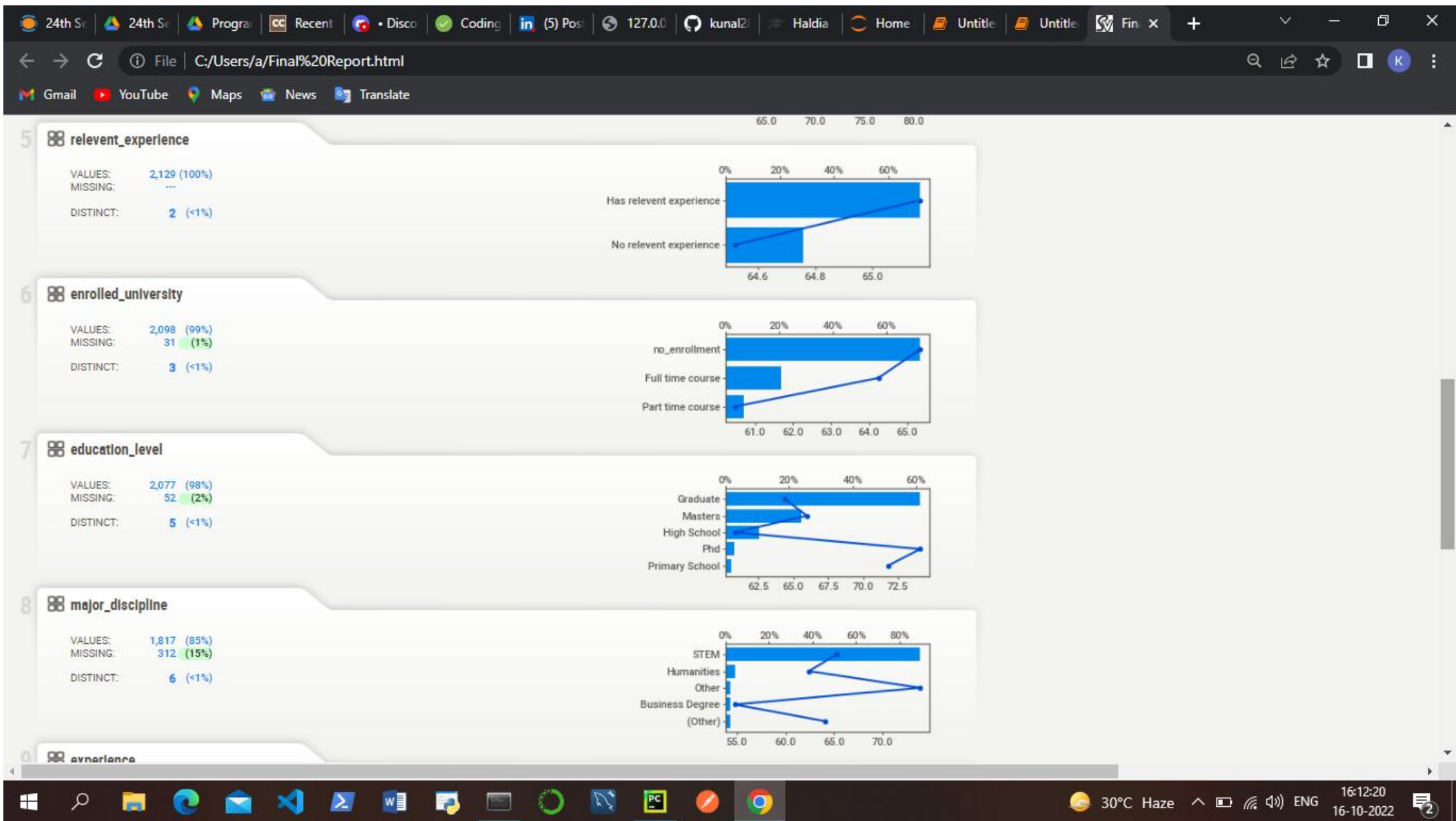
VALUES: 1,621 (76%)
MISSING: 508 (24%)
DISTINCT: 3 (<1%)

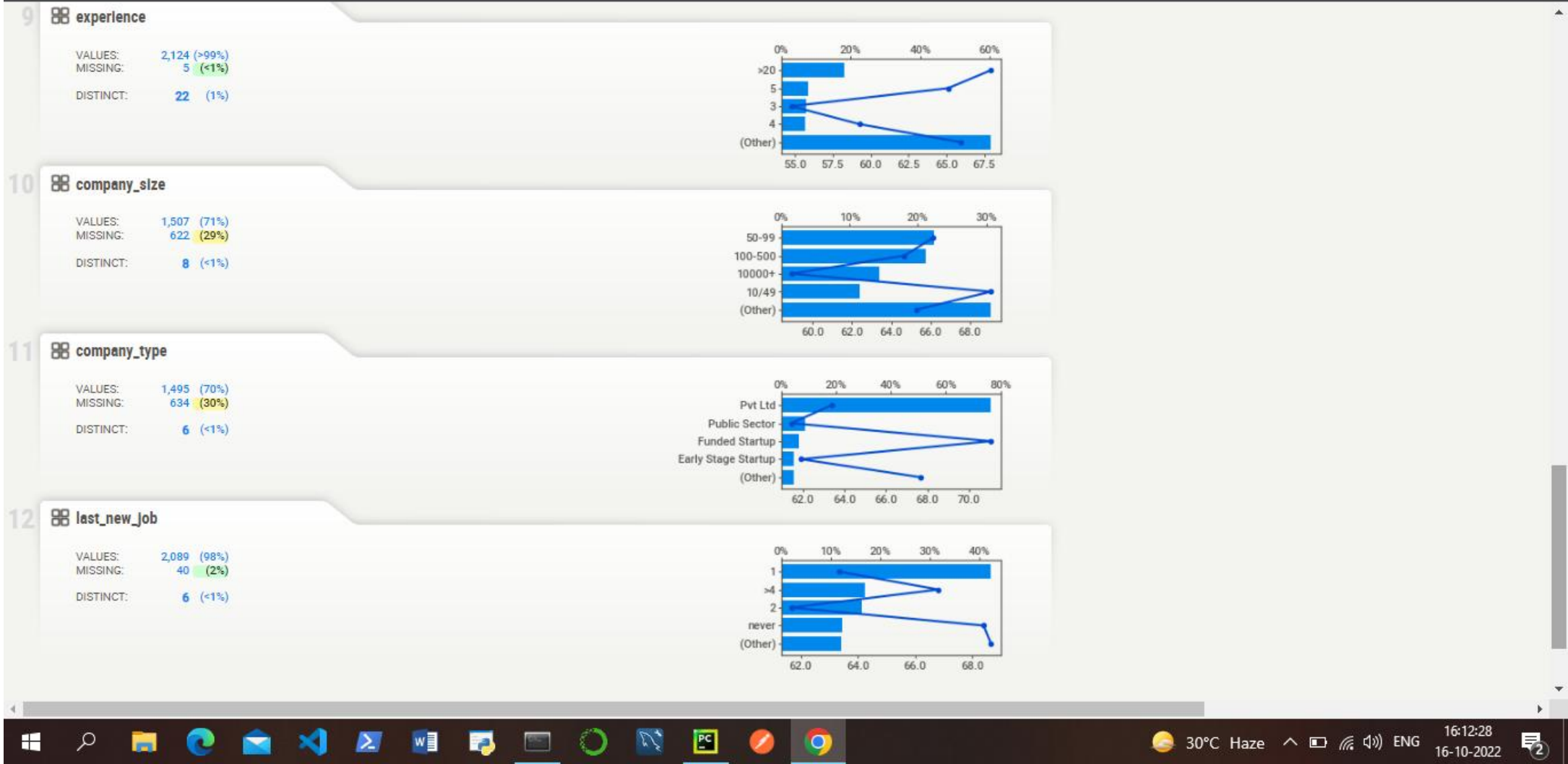


5 relevent_experience

VALUES: 2,129 (100%)
MISSING: ---
DISTINCT: 2 (<1%)







```
el-0.12.7 param-1.12.2 pyang-4.2.3 pyct-0.4.8 pyviz-comms-2.2.1 textblob-0.17.1 typing-extensions-4.4.0 wordcloud-1.8.2.2 xgb oost-1.6.2
```

```
In [4]: from autoviz.AutoViz_Class import AutoViz_Class
AV=AutoViz_Class()
df=AV.AutoViz(r'C:\Users\A\Downloads\Dataset\data0\aug_test.csv')
```

```
Shape of your Data Set loaded: (2129, 13)
#####
##### C L A S S I F Y I N G   V A R I A B L E S   #####
#####
Classifying variables in data set...
Data cleaning improvement suggestions. Complete them before proceeding to ML modeling.
```

	Nuniques	dtype	Nulls	Nullpercent	NuniquePercent	Value counts Min	Data cleaning improvement suggestions
enrollee_id	2129	int64	0	0.000000	100.000000	0	possible ID column: drop
training_hours	223	int64	0	0.000000	10.474401	0	
city	108	object	0	0.000000	5.072804	1	combine rare categories
city_development_index	81	float64	0	0.000000	3.804603	0	
experience	22	object	5	0.234852	1.033349	19	fill missing, fix mixed data types
company_size	8	object	622	29.215594	0.375763	68	fill missing, fix mixed data types

In []:

```
In [6]: from autoviz.AutoViz_Class import AutoViz_Class
        AV=AutoViz_Class()
```

```
In [10]: filename = "aug_test.csv"
        sep = ","
        dft = AV.AutoViz(
            filename,
            sep=",",
            depVar="",
            dfte=None,
            header=0,
            verbose=0,
            lowess=False,
            chart_format="svg",
            max_rows_analyzed=150000,
            max_cols_analyzed=30,
        )
```

```
pandas ascii encoder does not work for this file. Continuing...
pandas utf-8 encoder does not work for this file. Continuing...
pandas iso-8859-1 encoder does not work for this file. Continuing...
pandas cp1252 encoder does not work for this file. Continuing...
pandas latin1 encoder does not work for this file. Continuing...
Not able to read or load file. Please check your inputs and try again...
```

```
In [ ]:
```