

Ανάλυση του healthcare-dataset-stroke και γραφική αναπαράσταση του

Dataset statistics

Number of variables	12
Number of observations	5109
Missing cells	0
Missing cells (%)	0.0%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	479.1 KiB
Average record size in memory	96.0 B

Variable types

Numeric	4
Categorical	8

gender

Categorical

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	40.0 KiB

12994

02115

Toggle details

age

Real number ($\mathbb{R}_{\geq 0}$)

HIGH CORRELATION

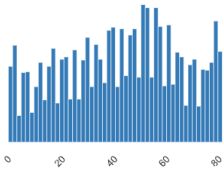
HIGH CORRELATION

HIGH CORRELATION

HIGH CORRELATION

Distinct	104
Distinct (%)	2.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	43.2299863

Minimum	0.08
Maximum	82
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	40.0 KiB



Toggle details

hypertension

Categorical

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	40.0 KiB

04611

1498

Toggle details

heart_disease

Categorical

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	40.0 KiB

04833

1276

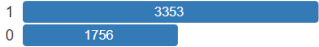
Toggle details

ever_married

Categorical

HIGH_CORRELATION
HIGH_CORRELATION
HIGH_CORRELATION
HIGH_CORRELATION
HIGH_CORRELATION

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	40.0 KiB



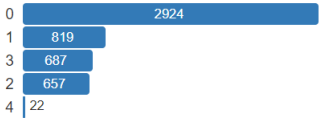
Toggle details

work_type

Categorical

HIGH_CORRELATION
HIGH_CORRELATION

Distinct	5
Distinct (%)	0.1%
Missing	0
Missing (%)	0.0%
Memory size	40.0 KiB

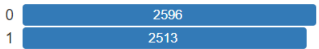


Toggle details

Residence_type

Categorical

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	40.0 KiB



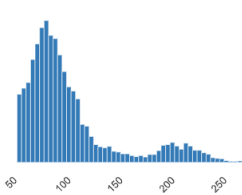
Toggle details

avg_glucose_level

Real number ($\mathbb{R}_{\geq 0}$)

Distinct	3978
Distinct (%)	77.9%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	106.1403993

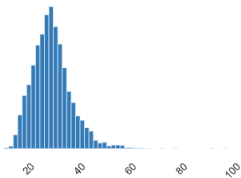
Minimum	55.12
Maximum	271.74
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	40.0 KiB



Toggle details

bmi
Real number ($\mathbb{R}_{\geq 0}$)
HIGH CORRELATION

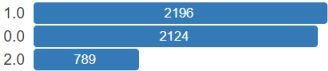
Distinct	601	Minimum	10.3
Distinct (%)	11.8%	Maximum	97.6
Missing	0	Zeros	0
Missing (%)	0.0%	Zeros (%)	0.0%
Infinite	0	Negative	0
Infinite (%)	0.0%	Negative (%)	0.0%
Mean	28.96097945	Memory size	40.0 KiB



Toggle details

smoking_status
Categorical

Distinct	3
Distinct (%)	0.1%
Missing	0
Missing (%)	0.0%
Memory size	40.0 KiB



Toggle details

stroke
Categorical

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	40.0 KiB



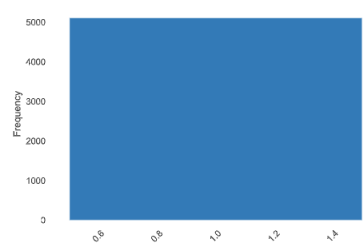
Toggle details

Overview Categories Words Characters

Common Values

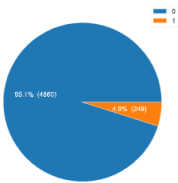
Value	Count	Frequency (%)
0	4860	95.1%
1	249	4.9%

Length

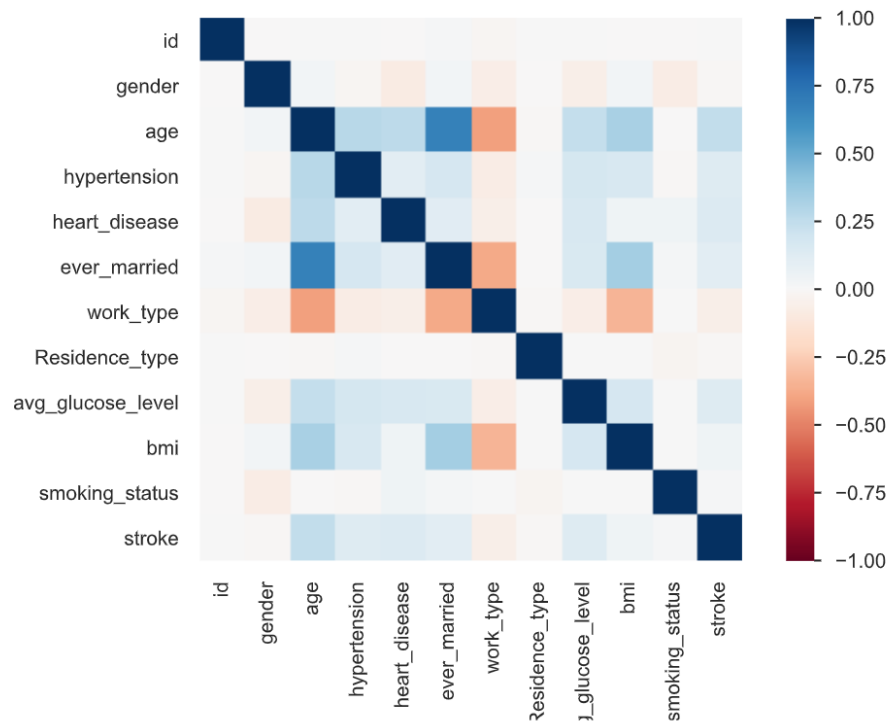


Histogram of lengths of the category

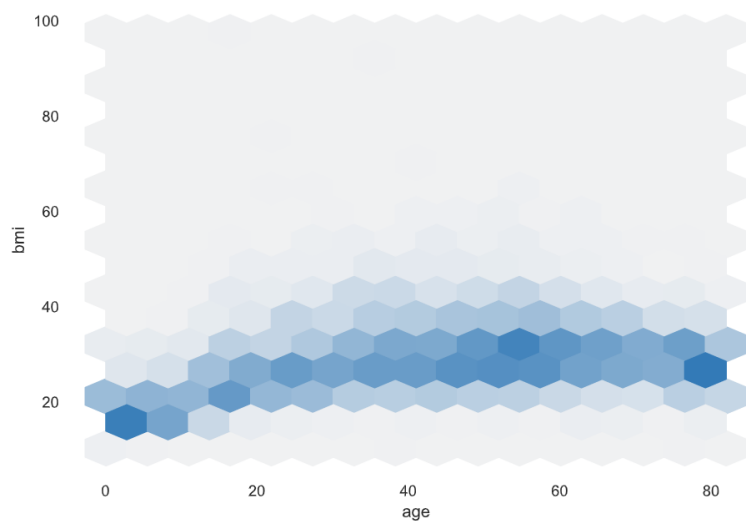
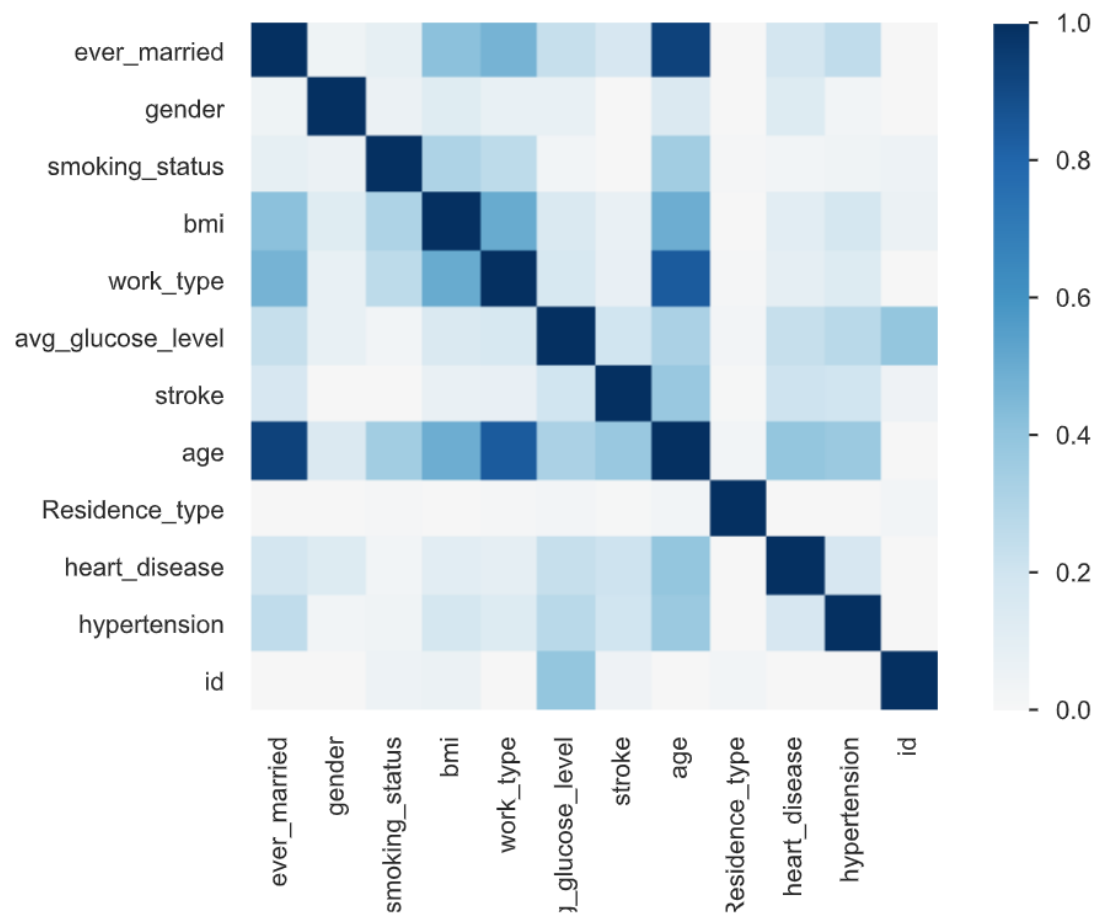
Pie chart



Coorelations
Pearson's



Φ_k Coorelation



age is highly correlated with ever_married

ever_married is highly correlated with age

age is highly correlated with ever_married

ever_married is highly correlated with age

age is highly correlated with ever_married

ever_married is highly correlated with age

ever_married is highly correlated with age

bmi is highly correlated with work_type

work_type is highly correlated with bmi and 1 other fields

age is highly correlated with ever_married and 1 other fields

ever_married is highly correlated with work_type

work_type is highly correlated with ever_married