

Manually Report About Stroke Data before Data Cleaning

=====

Number of Columns: 12

Number of Rows: 5110

Number of Duplicated Rows: 0

Total Number of Missing data: 201

Data Types

=====

id	int64
gender	object
age	float64
hypertension	int64
heart_disease	int64
ever_married	object
work_type	object
Residence_type	object
avg_glucose_level	float64
bmi	float64
smoking_status	object
stroke	int64

List of Numeric Variables

=====

id
age
hypertension
heart_disease
avg_glucose_level
bmi
stroke

List of Categorical Variables

=====

gender
ever_married
work_type
Residence_type
smoking_status

Number of NaN Values Per Column

=====

id	0
gender	0
age	0
hypertension	0
heart_disease	0
ever_married	0
work_type	0
Residence_type	0

avg_glucose_level	0
bmi	201
smoking_status	0
stroke	0

Number of Unique Values Per Column

```
=====
```

id	5110
gender	3
age	104
hypertension	2
heart_disease	2
ever_married	2
work_type	5
Residence_type	2
avg_glucose_level	3979
bmi	418
smoking_status	4
stroke	2

Basic Statistic About Numeric Variables

```
=====
```

Feature	mean	median	range	std_dev
id	36517.83	36932.00	72873.00	21161.72
age	43.23	45.00	81.92	22.61
hypertension	0.10	0.00	1.00	0.30
heart_disease	0.05	0.00	1.00	0.23
avg_glucose_level	106.15	91.88	216.62	45.28
bmi	28.89	28.10	87.30	7.85
stroke	0.05	0.00	1.00	0.22