GROUP MENBER:Yue Ma,Yuheng Li and Conger Yang;

5.

```python
import pandas as pd
#load the Dataset from the csv file
data = pd.read_csv("Stroke_dataset.csv")
```

6.

data   DataFrame   [5110, 12]   Column names: id, gender, age, hypertension, heart_disease, ever_marri ...

data - DataFrame

| 索引	id	gender	age	hypertension	heart_diseas	ever_married	work_type	residence_typ	_glucose_le	bmi
0	9046	Male	67	0	1	Yes	Private	Urban	228.69	36.6
1	51676	Female	61	0	0	Yes	Self-employed	Rural	202.21	nan
2	31112	Male	80	0	1	Yes	Private	Rural	105.92	32.5
3	60182	Female	49	0	0	Yes	Private	Urban	171.23	34.4
4	1665	Female	79	1	0	Yes	Self-employed	Rural	174.12	24
5	56669	Male	81	0	0	Yes	Private	Urban	186.21	29
6	53882	Male	74	1	1	Yes	Private	Rural	70.09	27.4
7	10434	Female	69	0	0	No	Private	Urban	94.39	22.8
8	27419	Female	59	0	0	Yes	Private	Rural	76.15	nan
9	60491	Female	78	0	0	Yes	Private	Urban	58.57	24.2
10	12109	Female	81	1	0	Yes	Private	Rural	80.43	29.7
11	12095	Female	61	0	1	Yes	Govt_job	Rural	120.46	36.8
12	12175	Female	54	0	0	Yes	Private	Urban	104.51	27.3
13	8213	Male	78	0	1	Yes	Private	Urban	219.84	nan

7.

```
In [2]: %runfile C:/Users/Acer/.spyder-py3/temp.py --wdir
      id  gender   age  ...   bmi  smoking_status stroke
0   9046    Male  67.0  ...  36.6  formerly smoked      1
1  51676  Female  61.0  ...   NaN     never smoked      1
2  31112    Male  80.0  ...  32.5     never smoked      1
3  60182  Female  49.0  ...  34.4           smokes      1
4   1665  Female  79.0  ...  24.0     never smoked      1

[5 rows x 12 columns]
```
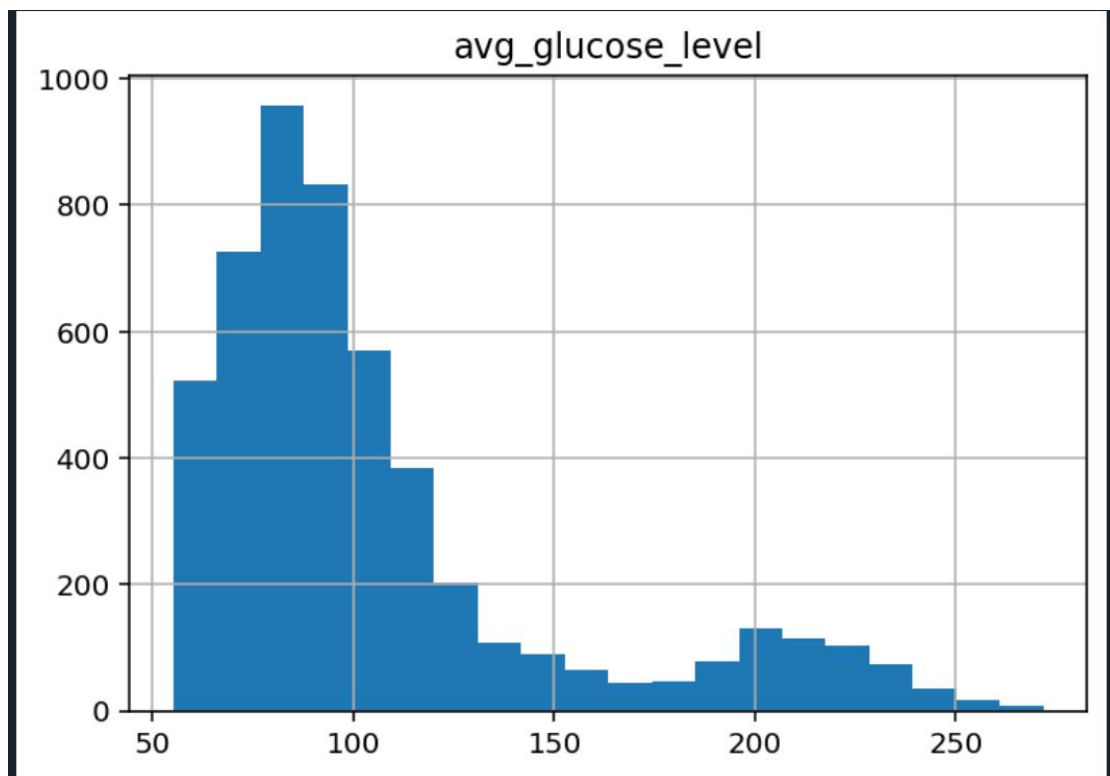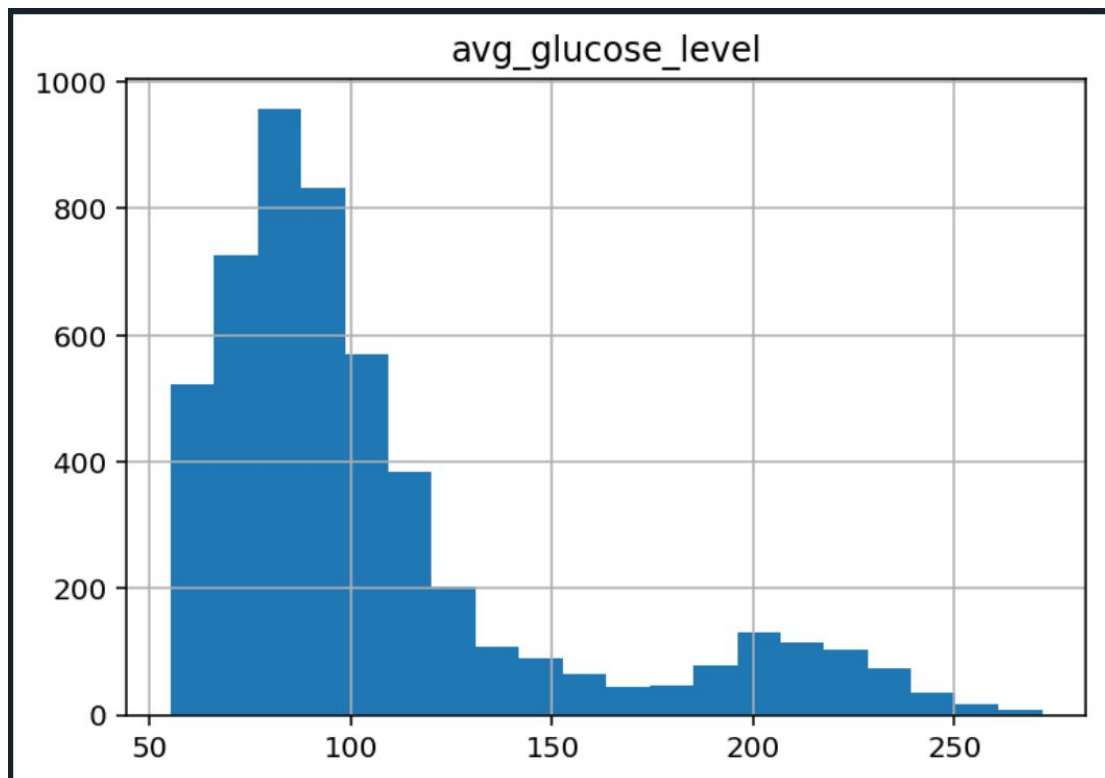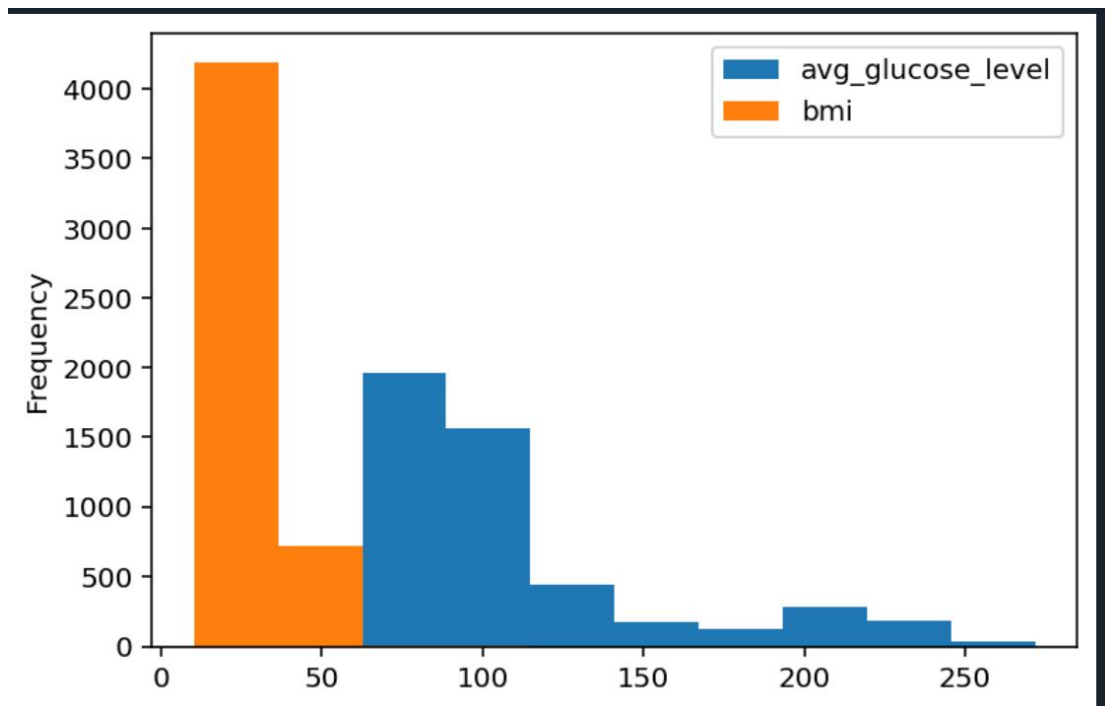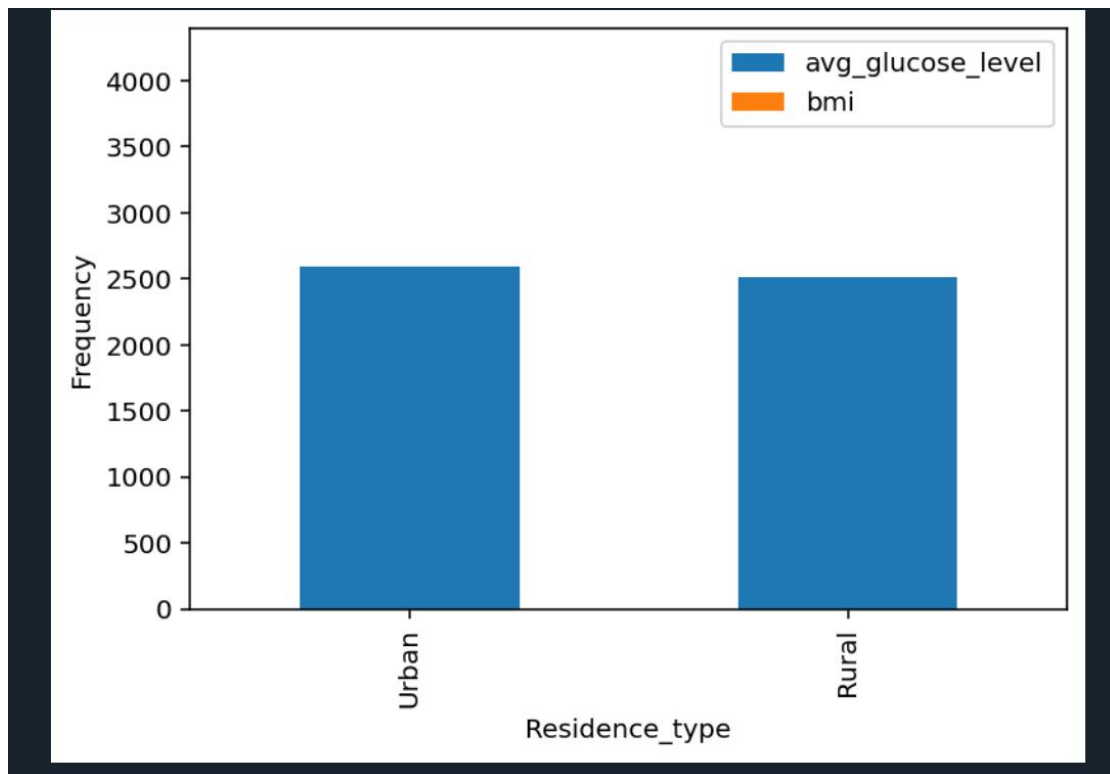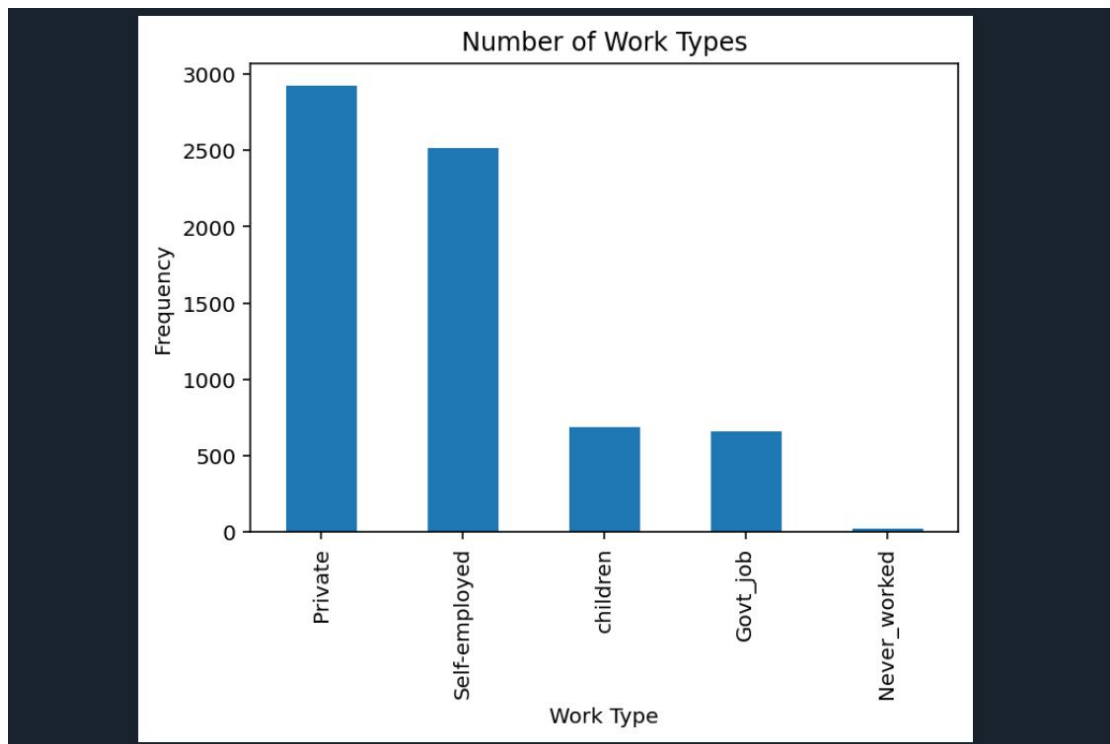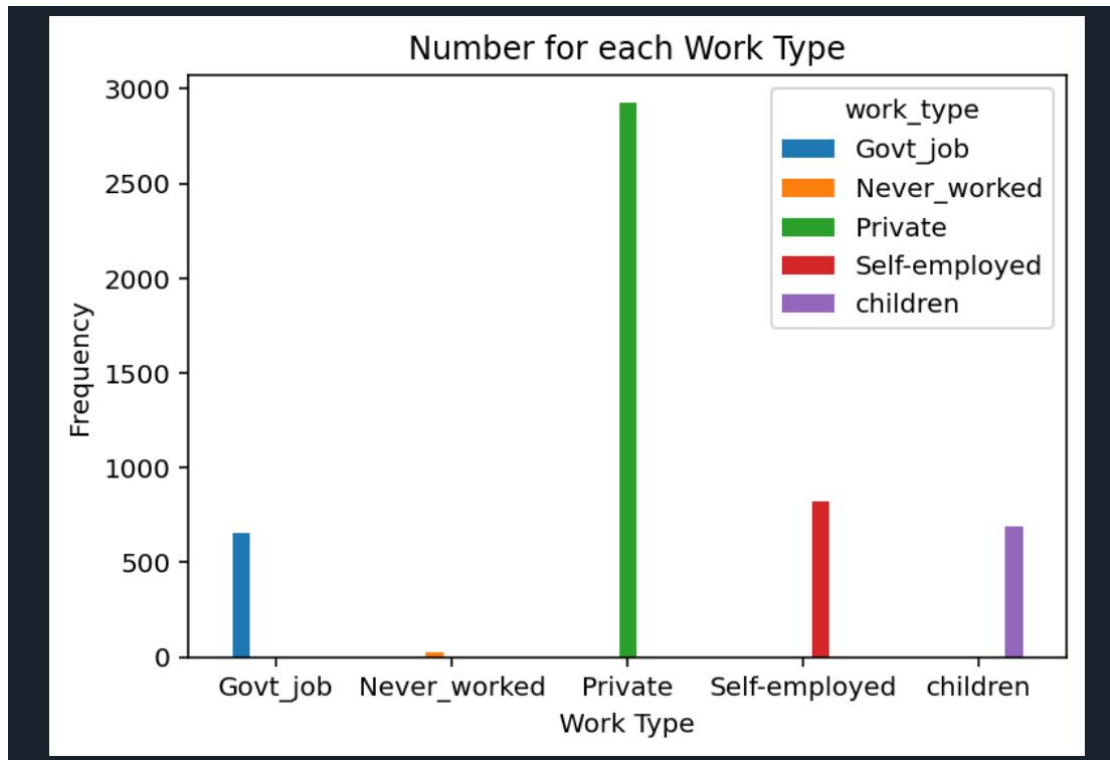
8.

9.

10.

avg_glucose_level
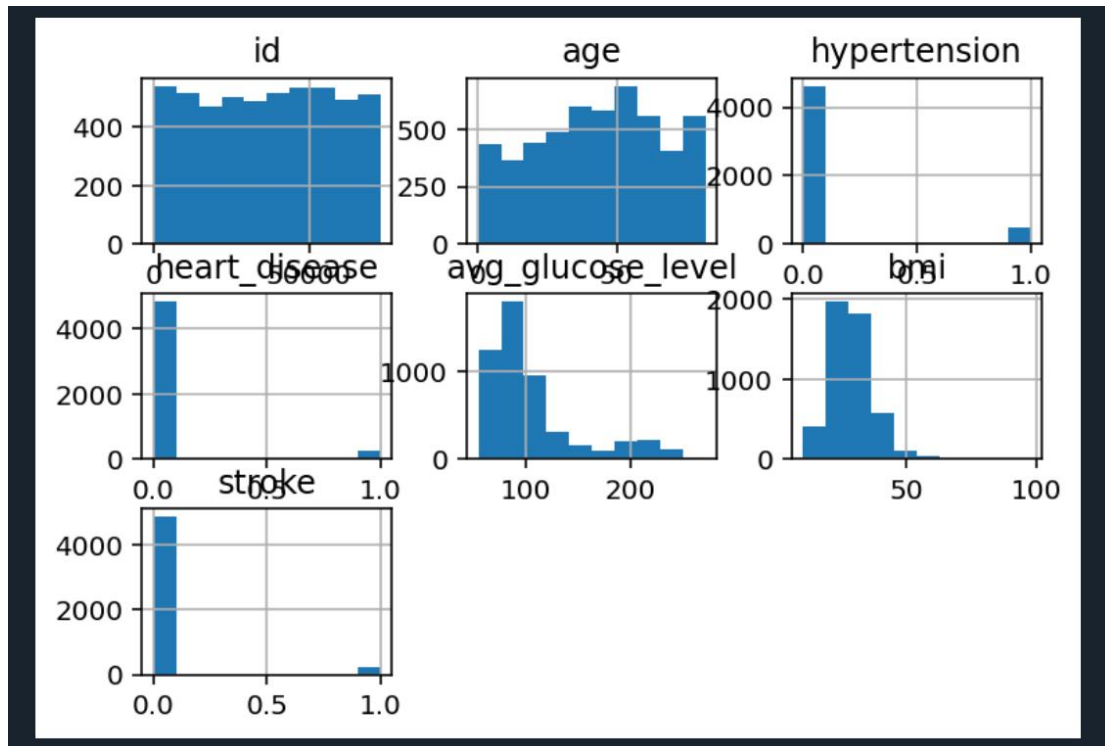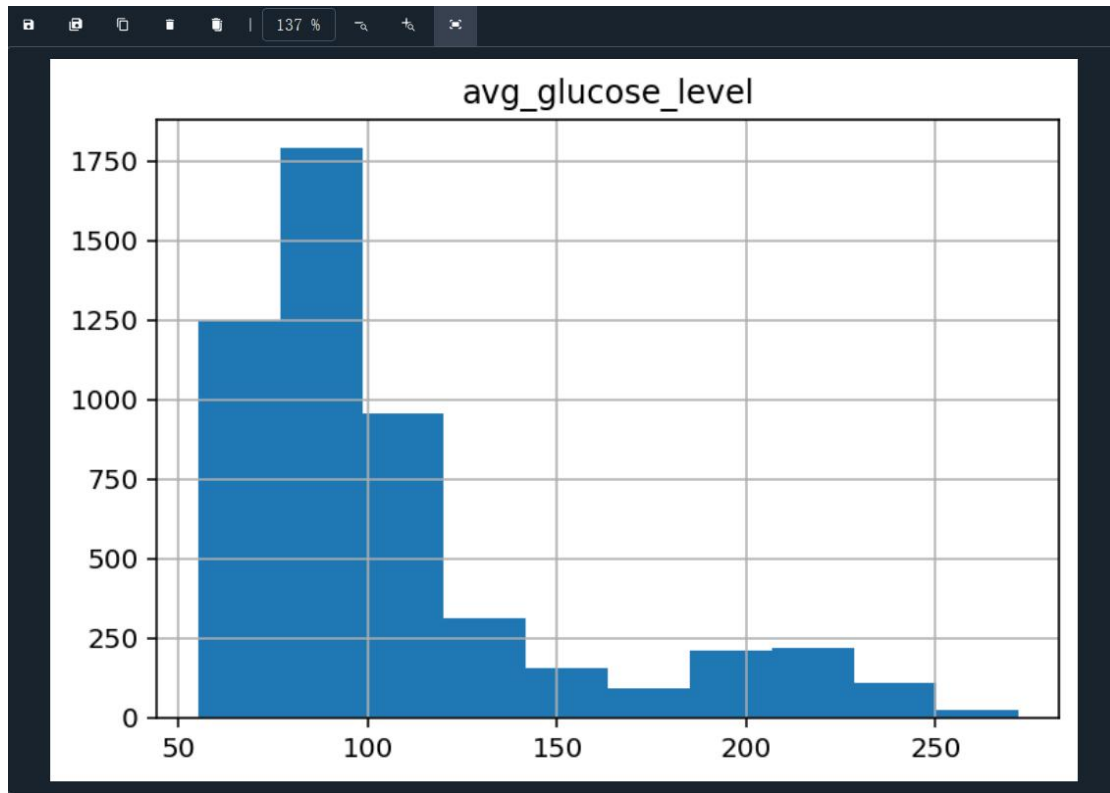
10.

11.

```
Residence_type
Urban    2596
Rural    2514
Name: count, dtype: int64
```
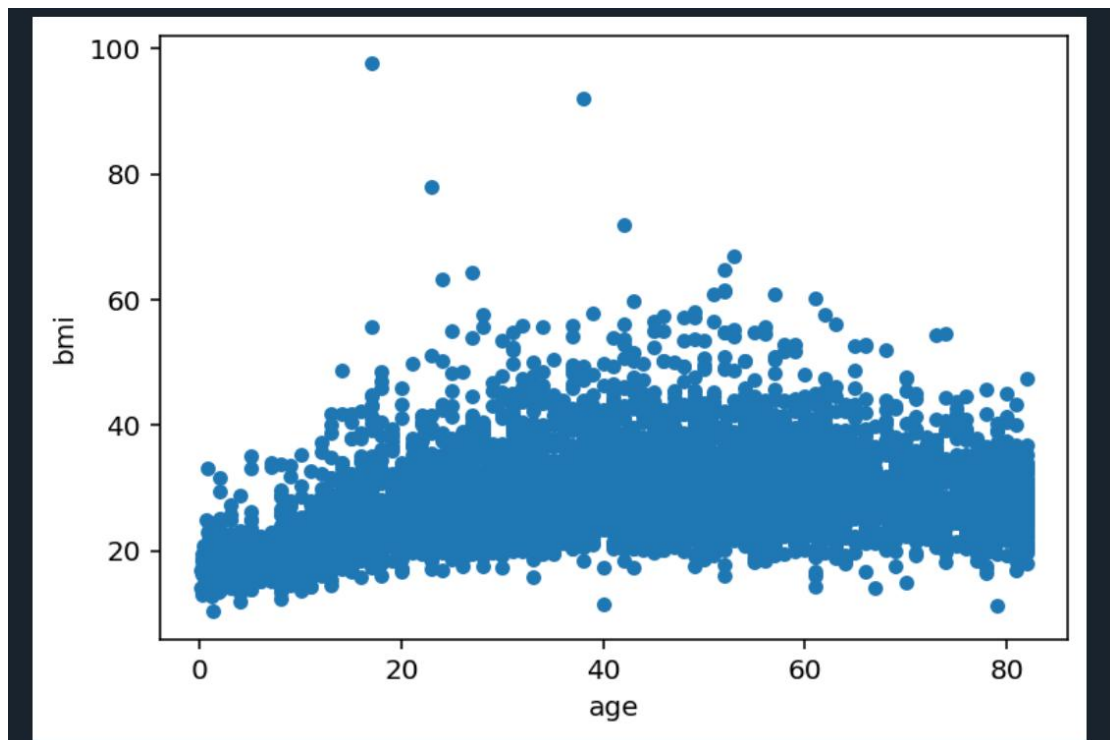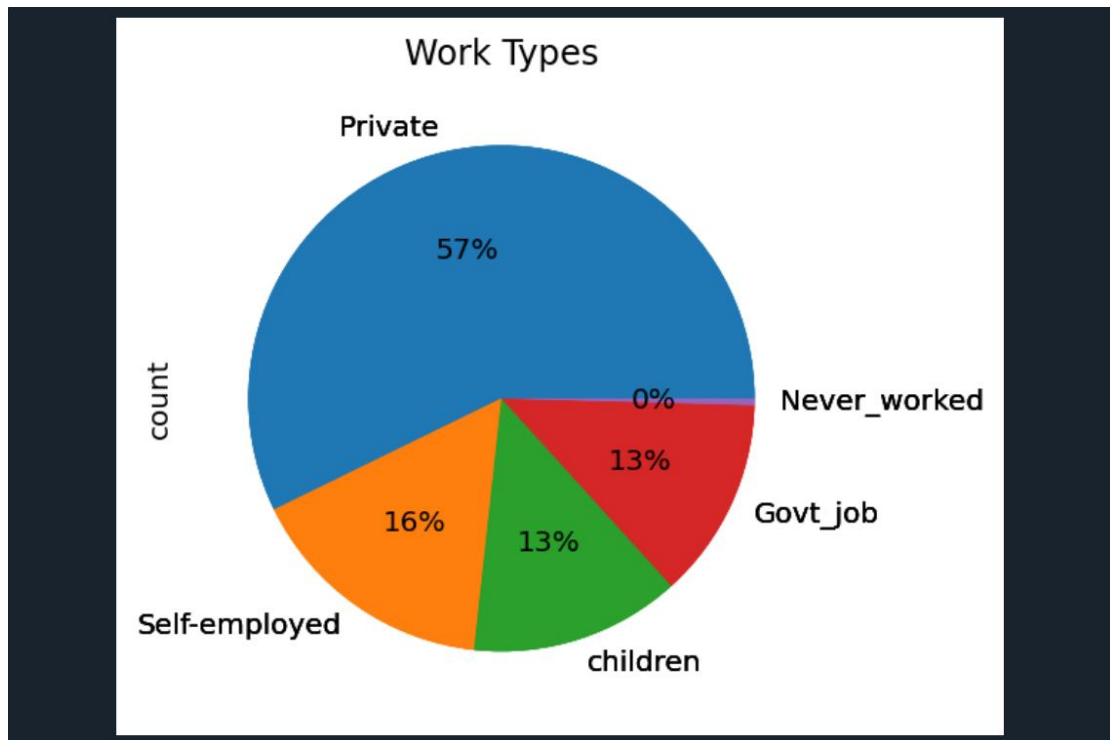
12.

13.

Number for each Work Type

14.

16.

17.