Dataset Overview:

The Diabetes Prediction Dataset includes patients' medical and demographic information as well as the state of their diabetes (positive or negative). The provided dataset contains various health-related attributes. The information includes gender, age, hypertension, heart disease, smoking history, BMI, HbA1c level, blood glucose level, and diabetes status. This dataset appears to be intended for research or analysis related to health and wellness, with a focus on factors that may impact an individual's health. The project's objective seems to be to investigate the relationships between these factors and their potential impact on developing diabetes. The findings from this analysis could be valuable for understanding how these factors are associated with specific health conditions and for making informed decisions regarding diabetes identification and personalized treatment.

<u>Reference:</u> https://www.kaggle.com/datasets/iammustafatz/diabetes-prediction-dataset/data

Data pre-processing:

*Importing the Dataset:

At first the dataset is imported from the correct path. The data is read from a CSV file named "Dataset_Midterm.csv" using the read.csv method in R. The resulting data frame is then assigned to a variable named mydata. mydata is then used to display the contents of the dataset in the console. The output of this code displays the entire contents of the dataset, including all rows and columns.

```
> options(max.print = 1e6)
> mydata<-read.csv("D:/Fall 2023-24 11th sem/Introduction to Data Science/Dataset_Midterm.csv", hea
der = TRUE, sep = ",")
> mydata
```

The output provides a way to verify that the data has been read correctly and allows for a quick visual inspection of the dataset. The output also provides information about the structure of the data, such as the number of columns and rows and the type of data in each column. Overall, this code and its output are a necessary first step in any data analysis project in R. The output is given below:

	gender	age	hypertension	heart_disease	smoking_history	bmi	Conso	le Te	rminal ×	Background Jobs ×			
1	Female	80	0	1	never	25.19				/ASUS/Downloads/ 🖈			
2	Female	54	0	0	No Info	27.32		ema i		1	0	never	32.02
3	Male	28	0	0	never	-27.32		emal emal		0	0	No Info No Info	29.30 27.32
4	Female	NA	0	0	current			emal:		0	0	never	24.93
5	Male	76	1	1				emal		0	0	never	19.95
6	Female	20	0	0	never	27.32		emal		0	0	never	18.03
7	I Cilia I C	79	0	0	No Info	23.86		emal		0	0	never	28.27
0	Mala	42	0	0				emal		Ö	Ö	No Info	
8	Male		0	0	never	33.64	46	Mal		0	Ō	never	27.32
9	Female	32	0	0	never	27.32	47 F	emal	e 43	0	0	No Info	27.32
10	Female	53	0	0	never		48 F	emal	e 74	0	0	No Info	28.12
11	Female	54	0	0	former	54.70	49 F	emal	e 21	0	0		26.10
12	Female	78	NA	0	former	36.05	50 F	emal		0	0	current	27.32
13	Female	67	0	0	never	25.69		emal		0	0	former	27.32
14	Female	76	0	0	No Info	27.32			e 290	0	0	not current	30.22
15		78	0	0	No Info	27.32		ema]		0	1	ever	23.11
16	Male	15	0	0	never	30.36		emal		0	0		27.32
17	Female	42	0	0	never	24.48		emal		0	0	No Info	28.16
18	Female	42	Ď.	0	No Info	27.32	56	Mal		0	0	never	26.78
19	Male	NA	0	0	ever	25.72	57	Mal		0	0	No Info	23.04
20	Male	40	0	0		36.38	58 59	Mal Mal		0	0	No Info No Info	15.94 15.80
			0	•	current			emal:		0	0	never	27.01
21	Male	5	0	0	No Info		61	Mal		0	0	never	27.32
22	Female	69	0	0	never	21.24		emal		0	0	never	22.19
23	Female	72	0	1	former			emal		0	0	never	27.45
24	Female	4	0	0	No Info	13.99		emal		Õ	ů.	No Info	
25	Male	30	0	0	never	33.76		emal		Õ	ő	never	26.45
26	Male	40	0	0	former	27.85	66	Mal		Ō	0	never	31.16
27	Male	45	NA	0	never	26.47	67	Mal	e 80	0	0	former	24.42
28	Male	43	0	0	never	26.08	68 F	emal	e 37	0	0	No Info	30.50
29	Female	53	0	0	No Info	31.75	69 I	emal	e 44	0	0	never	19.31
30	Male	50	0	0	No Info	25.15	70	Mal	e 67	0	1	not current	
31	Female	41	0	0	current		71	Mal		1	0	current	
32	Female	20	0	0		22.19	72	Mal		0	0	former	25.91
33		76	0	0	never	23.55		emal		0	0		27.32
	Female		0	•	never		74	Mal		0	0	former	37.16
34	Male	5	0	0	No Info	15.10		emal		0	0	never	63.48
35	Female	15	0	0	No Info	21.76	76	Mal		0	0	No Info	27.32
36	Female	26	0	0	never	21.22		emal		0	0	current	32.27
37	Male	5	0	0	No Info			emal emal		0	0	never never	27.32
20	Famala	77	1	1	novon	כח ככ	79 F	-ema I	e 0/		U	never	21.32

Cons	ala Tar	minal ×	Background Jobs ×			
			_			
R			'ASUS/Downloads/ 🖈			
80	Female		1	0	never	27.32
81	Female		0	0	never	31.70
82	Male		0	0	never	22.06
83	Female		0	0	never	36.49
84	Male		0	0	current	30.80
85	Female		0	0	not current	39.36
86	Male		0	0	never	31.90
87	Female	43	0	0	never	26.71
88	Male		0	0	No Info	27.32
89	Male	57	1	1	not current	27.77
90	Female	43	0	0	never	27.32
91	Male	63	1	0	ever	35.06
92	Male	80	0	0	never	23.25
93	Female	70	0	0	current	29.25
94	Female	42	0	0	never	24.81
95	Female	80	0	0	former	36.18
96	Female	52	1	0	never	50.30
97	Male	71	0	0	never	27.09
98	Female	43	0	0	never	27.32
99	Male	71	0	0	never	27.09
100	Male	80	0	1	former	24.36
101	Male	59	0	0	current	29.20
102	Male	29	0	0	current	25.41
103	Female	68	0	0	No Info	40.31
104	Female	52	0	0	No Info	27.32
105	Male	71	0	0	never	26.53
106	Male	48	1	0	current	36.12
107	Female	79	1	0	former	27.32
108	Male	37	0	0	never	37.24
109	Female	73	0	0	never	35.56
110	Female	59	0	0	former	43.41
111	Female	80	0	0	never	27.32
112	Female	64	0	0	ever	49.27
113	Male		0	0	never	39.00
114	Male		Ö	Ö	never	22.43
115	Male		Ö	ŏ	not current	32.19
116	Male		í	ő	ever	25.94
	Female		0	ŏ	never	27.73
118	Male		ő	ŏ	ever	19.46
	Female		0	ő	No Info	27.32
	Female		0	0	No Info	27.32
120	· Cina i C	. 45	Ü	V	110 1110	21.32

Con	sole	Terminal ×	Background Jobs ×	
R	R 4.	3.1 · C:/Users/A	SUS/Downloads/ 🖈	
	HbA	TC_level I	blood_glucose_level	diabetes
1		6.6	140	0
2		6.6	80	0
3		5.7	158	0
4		5.0	155	0
5		4.8	155	0
6		6.6	85	0
7		5.7	85	0
8		4.8	145	0
9		5.0	100	0
10		6.1	85	0
11		6.0	100	0
12		5.0	130	0
13		5.8	200	0
14		5.0	160	0
15		6.6	126	0
16		6.1	200	0
17		5.7	158	0
18		5.7	80	0
19		3.5	159	0
20		6.0	90	0
21		6.2	85	0
22		4.8	85	0
23		6.5	130	0
24		4.0	140	0
25		6.1	126	0
26		5.8	80	0
27		4.0	158	0
28		6.1	155	0
29		4.0	200	0
30		4.0	145	0
31		6.2	126	0
32		3.5	100	0
33		5.0	85	0
34		5.8	85	0
35		4.5	130	0
36		6.6	200	0
37		6.6	130	0
38		5.0	159	0
39		4.8	159	0
40		3.5	160	0
41		6 1	100	0

```
155
155
        Terminal × Background Jobs
                                                                       83
    R 4.3.1 · C:/Users/ASUS/Downloads/
                                                                                      6.6
                                                                                                                280
42
              5.0
                                                                                                                140
                                                                       86
                                                                                                                155
44
45
46
47
48
49
              6.2
6.5
                                                                                      6.5
                                                                       87
                                                                                                                300
                                                                       88
                                                                                      6.8
                                                                                                                159
                                     155
                                                                       89
                                                                                      6.6
                                                                                                                160
                                                                                                                155
                                     100
                                     158
159
50
51
52
53
54
55
56
67
58
60
61
62
63
64
65
67
77
77
77
77
77
80
                                                                       92
                                                                                                                159
                                                                       93
                                                                                      8.2
                                                                                                                130
                                                                       94
                                                                                      9.0
                                                                                                                159
                                                                       95
                                                                                                                200
                                                                                      6.5
                                     145
              5.0
                                                                       97
                                                                                                                200
                                     200
                                                                       98
                                                                                      6.6
                                                                                                                130
                                     160
                                                                       99
                                                                                                                200
                                                                       100
                                                                                                                280
                                                                       101
                                                                                                                220
                                     126
                                                                       102
                                                                                      6.1
                                                                       103
                                                                                                                260
                                     130
                                                                                      9.0
                                                                       104
                                                                                                                140
                                                                       105
                                                                                                                159
                                                                                      8.8
                                     158
                                                                       106
                                                                                                                140
                                                                                      6.8
                                                                       107
                                                                                      6.5
                                                                                                                159
                                     160
                                                                       108
                                                                                                                126
                                     200
                                                                       109
                                                                                                                260
                                                                       110
                                                                                      6 2
                                                                                                                160
                                                                                                                200
                                                                                      6.0
                                                                       111
                                                                                                                140
                                                                       112
                                                                                      8.2
                                     159
159
                                                                       113
                                                                       114
                                                                                                                160
                                                                       115
                                                                                      5.8
                                                                                                                300
                                                                       116
                                                                                      9.0
                                                                                                                140
                                                                                                                145
                                                                       117
                                                                                      8.8
                                     300
                                                                       118
                                                                                      9.0
                                                                                                                130
              6.8
                                     280
                                                                       120
                                                                                                                159
> names (mydata)
[1] "gender"
                                        "age"
                                                                            "hypertension"
                                                                                                               "heart_disease"
     "smoking_history"
                                         "bmi"
                                                                            "HbA1c_level"
                                                                                                               "blood_glucose_level"
[9] "diabetes"
```

* Identifying Missing Values:

To identify missing value "is.na" function has been used to check for missing values in the variable. The is.na function returns a logical value of TRUE for each missing value in the data and FALSE for all other values.

The output of this code is a matrix that has the same dimensions as the variable. The matrix contains TRUE values in the cells where the corresponding values are missing and FALSE values in all other cells. This output is useful for identifying any missing values in the dataset and determining the extent to which missing values may affect subsequent analysis:

		und Jobs ×				Collsole	Terminal × Backgro	und Jobs ×			
R 4.	3.1 · C:/Users/ASUS/Down genger age nvi		rt_aisease smok	ng pistom/ pmi up	ATC_level	R 4.	3.1 - C:/Users/ASUS/Dowr	loads/ 🔎			
1 1	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE	[83,]	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
2,]	FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE	[84,]	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
3.1	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE	[85.]	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
4,]	FALSE TRUE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
5.1	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
5,] 6.]	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE
7.1	TRUE FALSE	FALSE	FALSE	TRUE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
8,]	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
9.1	FALSE FALSE		FALSE		FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
	FALSE FALSE	FALSE FALSE	FALSE	FALSE FALSE FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
0,] 1.]	FALSE FALSE		FALSE		FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
2.1	FALSE FALSE	FALSE TRUE	FALSE	FALSE FALSE FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
3.1	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
4.1							FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
5.1	FALSE FALSE TRUE FALSE	FALSE FALSE	FALSE FALSE	TRUE FALSE TRUE FALSE	FALSE FALSE		FALSE FALSE		FALSE	FALSE FALSE	FALSE
								FALSE			
5,]	FALSE FALSE FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
,]	FALSE FALSE	FALSE FALSE	FALSE FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
,]	FALSE FALSE			TRUE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
,,]	FALSE TRUE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
),]]	FALSE FALSE	FALSE FALSE	FALSE FALSE	FALSE FALSE TRUE FALSE	FALSE FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
2,]	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE
3.1		FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE
4.1	FALSE FALSE FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
5.1	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
5,]	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
.;	FALSE FALSE	TRUE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
::	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
9.1	FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
0.1	FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
.;	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
.;	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
,, j i, j	FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
.;	FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE		FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
:1	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE	[117,]	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
.;	FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE	[118,]	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE
3,]	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE	[119,]	FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE
	FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE	[120.]	FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE
),]),]	FALSE FALSE	FALSE	FALSE	TRUE FALSE	FALSE		blood_glucose_le				
	FALSE FALSE	FALSE	FALSE	FALSE FALSE	FALSE FALSE	[1,]		LSE FALSE			
	FALSE PALSE	FALSE	FALSE	FALSE FALSE	FALSE.	[2,]		LSE FALSE			

onsole	Terminal ×	Background Joi	os ×				
R R4.3	B.1 · C:/Users/A	SUS/Downloads/					
[42,]	FALSE FAI			FALSE	FALSE F	ALSE	FALSE
[43,]	FALSE FAI	LSE I	FALSE	FALSE	FALSE F	ALSE	FALSE
[44,]	FALSE FAI	LSE I	FALSE	FALSE	FALSE F	ALSE	FALSE
[45,]	FALSE FA	LSE I	FALSE	FALSE	TRUE F	ALSE	FALSE
[46,]	FALSE FAI	LSE I	FALSE	FALSE	FALSE F	ALSE	FALSE
[47,]	FALSE FA	LSE I	FALSE	FALSE	TRUE F	ALSE	FALSE
[48,]	FALSE FAI	LSE I	FALSE	FALSE	TRUE F	ALSE	FALSE
[49,]	FALSE FAI	LSE I	FALSE	FALSE	TRUE F	ALSE	FALSE
[50,]	FALSE FAI	LSE I	FALSE	FALSE	FALSE F	ALSE	FALSE
[51,]	FALSE FAI	LSE I	FALSE	FALSE	FALSE F	ALSE	FALSE
[52,]	FALSE FAI	LSE I	FALSE	FALSE	FALSE F	ALSE	FALSE
[53,]	FALSE FAI			FALSE	FALSE F		FALSE
[54,]	FALSE FA	LSE I	FALSE	FALSE	TRUE F	ALSE	FALSE
[55,]	FALSE TI	RUE I	FALSE	FALSE	TRUE F	ALSE	FALSE
[56,]	FALSE FAI	LSE I	FALSE	FALSE	FALSE F	ALSE	FALSE
[57,]	FALSE FAI	LSE I	FALSE	FALSE	TRUE F	ALSE	FALSE
[58,]	FALSE FAI	LSE I	FALSE	FALSE	TRUE F	ALSE	FALSE
[59,]	FALSE FAI	LSE I	FALSE	FALSE	TRUE F	ALSE	FALSE
60.]	FALSE FAI	LSE I	FALSE	FALSE	FALSE F	ALSE	FALSE
61.]	FALSE FAI	LSE	FALSE	FALSE	TRUE F	ALSE	FALSE
[62,]	FALSE FAI	LSE I	FALSE	FALSE	FALSE F	ALSE	FALSE
[63,]	FALSE FAI			FALSE	FALSE F		FALSE
[64,]	FALSE FAI	LSE I	FALSE	FALSE	TRUE F	ALSE	FALSE
[65,]	FALSE FAI			FALSE	FALSE F		FALSE
[66,]	FALSE FAI	LSE I	FALSE	FALSE	FALSE F	ALSE	FALSE
[67,]	FALSE FAI	LSE I	FALSE	FALSE	FALSE F	ALSE	FALSE
[68.]	FALSE FAI			FALSE	TRUE F		FALSE
[69,]	FALSE FAI	LSE I	FALSE	FALSE	FALSE F	ALSE	FALSE
[70,]	FALSE FAI			FALSE	FALSE F		FALSE
[71,]	FALSE FAI			FALSE	FALSE F		FALSE
72,]	FALSE FAI			FALSE	FALSE F		FALSE
[73,]	FALSE FAI	LSE	FALSE	FALSE	FALSE F	ALSE	FALSE
[74.]	FALSE FA			FALSE	FALSE F		FALSE
[75,]	FALSE FAI			FALSE	FALSE F		FALSE
[76,]	FALSE FAI			FALSE	TRUE F		FALSE
77,]	FALSE FAI			FALSE	FALSE F		FALSE
[78.]	FALSE FAI			FALSE	FALSE F		FALSE
[79,]	FALSE FAI			FALSE	FALSE F		FALSE
[80,]	FALSE FAI			FALSE	FALSE F		FALSE
[81.]	FALSE TI			FALSE	FALSE F		FALSE
[82,]	FALSE FAI			FALSE	FALSE F		FALSE

Console	Terminal ×	Background Jobs	<	Con
R 4.3	3.1 · C:/Users/A	ASUS/Downloads/ 🗇		Q
[4,J		FALSE	FALSE	
[5,]		FALSE	FALSE	[4
[6,]		FALSE	FALSE	[4
[7,]		FALSE	FALSE	[4
[8,]		FALSE	FALSE	[5
[9,]		FALSE	FALSE	[5
[10,]		FALSE	FALSE	[5
[11,]		FALSE	FALSE	[5
[12,]		FALSE	FALSE	[5
[13,]		FALSE	FALSE	[5
[14,]		FALSE	FALSE	[5
[15,]		FALSE	FALSE	
[16,]		FALSE	FALSE	[5
[17,]		FALSE	FALSE	[5
[18,]		FALSE	FALSE	[5
[19,]		FALSE	FALSE	[6
[20,]		FALSE	FALSE	[6
[21,]		FALSE	FALSE	[6
[22,]		FALSE	FALSE	[6
[23,]		FALSE	FALSE	[6
[24,]		FALSE	FALSE	[6
[25,]		FALSE	FALSE	[6
[26,]		FALSE	FALSE	[6
[27,]		FALSE	FALSE	[6
[28,]		FALSE	FALSE	[6
[29,]		FALSE	FALSE	[7
[30,]		FALSE	FALSE	[7
[31,]		FALSE	FALSE	[7
[32,]		FALSE	FALSE	[7
[33,]		FALSE	FALSE	[7
[34,]		FALSE	FALSE	[7
[35,]		FALSE	FALSE	[7
[36,]		FALSE	FALSE	[7
[37,]		FALSE	FALSE	[7
[38,]		FALSE	FALSE	[7
[39,]		FALSE	FALSE	[8
[40,]		FALSE	FALSE	[8
[41,]		FALSE	FALSE	[8
[42,]		FALSE	FALSE	[8]
[43,]		FALSE	FALSE	[8]
[44,]		FALSE	FALSE	[8
Γ45.1		FALSE	FALSE	

R 4.3.1 · C:/Users/ASUS/Downloads/
[46,] FALSE FALSE [47,] FALSE FALSE [48,] FALSE FALSE [49,] FALSE FALSE [50,] FALSE FALSE [51,] FALSE FALSE [52,] FALSE FALSE [53,] FALSE FALSE [54,] FALSE FALSE [55,] FALSE FALSE [56,] FALSE FALSE
[47,] FALSE FALSE [48,] FALSE FALSE [49,] FALSE FALSE [50,] FALSE FALSE [51,] FALSE FALSE [52,] FALSE FALSE [53,] FALSE FALSE [54,] FALSE FALSE [55,] FALSE FALSE [56,] FALSE FALSE [56,] FALSE FALSE
[48,] FALSE FALSE [49,] FALSE FALSE [50,] FALSE FALSE [51,] FALSE FALSE [52,] FALSE FALSE [53,] FALSE FALSE [54,] FALSE FALSE [55,] FALSE FALSE [56,] FALSE FALSE
[49,] FALSE FALSE [50,] FALSE FALSE [51,] FALSE FALSE [52,] FALSE FALSE [53,] FALSE FALSE [54,] FALSE FALSE [55,] FALSE FALSE [56,] FALSE FALSE
[50,] FALSE FALSE [51,] FALSE FALSE [52,] FALSE FALSE [53,] FALSE FALSE [54,] FALSE FALSE [55,] FALSE FALSE [56,] FALSE FALSE
[51,] FALSE FALSE [52,] FALSE FALSE [53,] FALSE FALSE [54,] FALSE FALSE [55,] FALSE FALSE [56,] FALSE FALSE
[52,] FALSE FALSE [53,] FALSE FALSE [54,] FALSE FALSE [55,] FALSE FALSE [56,] FALSE FALSE
[53,] FALSE FALSE [54,] FALSE FALSE [55,] FALSE FALSE [56,] FALSE FALSE
[54,] FALSE FALSE [55,] FALSE FALSE [56,] FALSE FALSE
[55,] FALSE FALSE [56,] FALSE FALSE
[56,] FALSE FALSE
[5/,] FALSE FALSE
[58,] FALSE FALSE
[59,] FALSE FALSE
[60,] FALSE FALSE
[61,] FALSE FALSE
[62,] FALSE FALSE
[63,] FALSE FALSE
[64,] FALSE FALSE
[65,] FALSE FALSE
[66,] FALSE FALSE
[67,] FALSE FALSE
[68,] FALSE FALSE
[69,] FALSE FALSE
[70,] FALSE FALSE
[71,] FALSE FALSE
[72,] FALSE FALSE [73,] FALSE FALSE
[74,] FALSE FALSE [75,] FALSE FALSE
[76,] FALSE FALSE
[77,] FALSE FALSE
[78,] FALSE FALSE
[79,] FALSE FALSE
[80,] FALSE FALSE
[81,] FALSE FALSE
[82,] FALSE FALSE
[83,] FALSE FALSE
[84,] FALSE FALSE
[85,] FALSE FALSE
[86.] FALSE FALSE

[87,]	FALSE	FALSE
[88,]	FALSE	FALSE
[89,]	FALSE	FALSE
[90,]	FALSE	FALSE
[91,]	FALSE	FALSE
[92,]	FALSE	FALSE
[93,]	FALSE	FALSE
[94,]	FALSE	FALSE
[95,]	FALSE	FALSE
[96,]	FALSE	FALSE
[97,]	FALSE	FALSE
[98,]	FALSE	FALSE
[99,]	FALSE	FALSE
[100,]	FALSE	FALSE
[101,]	FALSE	FALSE
[102,]	FALSE	FALSE
[103,]	FALSE	FALSE
[104,]	FALSE	FALSE
[105,]	FALSE	FALSE
[106,]	FALSE	FALSE
[107,]	FALSE	FALSE
[108,]	FALSE	FALSE
[109,]	FALSE	FALSE
[110,]	FALSE	FALSE
[111,]	FALSE	FALSE
[112,]	FALSE	FALSE
[113,]	FALSE	FALSE
[114,]	FALSE	FALSE
[115,]	FALSE	FALSE
[116,]	FALSE	FALSE
[117,]	FALSE	FALSE
[118,]	FALSE	FALSE
[119,]	FALSE	FALSE
[120,]	FALSE	FALSE
>		

The mydata Output after replacing missing values with NA:

								Cons	sole Tern	ninal ×	Background Jobs ×				
> 1	ıydata							R	R 4.3.1 · 0	:/Users/	/ASUS/Downloads/ 🖈				
				heart_disease	smoking_history		HbA1c_level	39	⊦ema re		Ú	U	<na></na>	29.30	4.8
1	Female		0	1		25.19	6.6	40	Female	67	0	0	<na></na>	27.32	3.5
2	Female		0	0	<na></na>	27.32	6.6	41	Female	44	0	0	never	24.93	6.1
3	Male	28	0	0		-27.32	5.7	42	Female	29	0	0	never	19.95	5.0
4	Female	NA	0	0	current		5.0	43	Female	60	0	0	never	18.03	4.0
5	Male	76	1	1	current	20.14	4.8	44	Female	38	0	0	never	28.27	6.2
6	Female	20	0	0	never	27.32	6.6	45	Female	3	0	0	<na></na>	19.27	6.5
7	<na></na>	79	0	0	<na></na>	23.86	5.7	46	Male		0	0	never	27.32	6.1
8	Male	42	0	0	never	33.64	4.8	47	Female	43	0	0	<na></na>	27.32	4.0
9	Female	32	0	0	never	27.32	5.0	48	Female		0	0	<na></na>	28.12	5.0
10	Female	53	0	0	never	27.32	6.1	49	Female		0	0	<na></na>	26.10	5.8
11	Female	54	0	0	former	54.70	6.0	50	Female		0	0	current	27.32	6.5
12	Female	78	NA	0	former	36.05	5.0	51	Female		0	0	former	27.32	6.0
13	Female	67	0	0	never	25.69	5.8	52	Female		0	0	not current	30.22	5.7
14	Female	76	0	0	<na></na>	27.32	5.0	53	Female		0	1	ever	23.11	6.5
15	<na></na>	78	0	0	<na></na>	27.32	6.6	54	Female		0	0	<na></na>	27.32	5.7
16	Male	15	0	0	never	30.36	6.1	55	Female		0	0	<na></na>	28.16	5.0
17	Female	42	0	0	never	24.48	5.7	56	Male		0	0	never	26.78	4.8
18	Female	42	0	0	<na></na>	27.32	5.7	57	Male		0	0	<na></na>	23.04	5.7
19	Male	NA	0	0	ever	25.72	3.5	58	Male	7	0	0	<na></na>	15.94	5.8
20	Male	40	0	0	current	36.38	6.0	59	Male	3	0	0		15.80	6.2
21	Male	5	0	0	<na></na>	18.80	6.2	60	Female		0	0	never	27.01	6.2
22		69	0	0	never	21.24	4.8	61	Male		0	0	<na></na>	27.32	3.5
23	Female	72	0	1	former	27.94	6.5	62	Female		0	0	never	22.19	6.6
24	Female	4	Ö	ō		13.99	4.0	63 64	Female		0	0	never	27.45 17.98	5.7 6.5
25	Male	30	Ö	ŏ	never	33.76	6.1	65	Female Female		0	0		26.45	5.7
26	Male	40	0	ŏ	former	27.85	5.8	66	Male		0	0	never never	31.16	5.8
27	Male		NA	ŏ	never	26.47	4.0	67	Male		0	0	former	24.42	4.0
28		43		Ŏ	never	26.08	6.1	68	Female		0	0	<na></na>	30.50	5.7
29		53	0	0	<na></na>	31.75	4.0	69	Female		0	0	never	19.31	6.5
30		50	0	0	<na></na>	25.15	4.0	70	Male		0	1			6.5
31		41	0	0	current	22.01	6.2	71	Male		1	0	current		5.7
32	Female	20	0	0	never	22.19	3.5	72	Male		0	0	former		9.0
33		76	0	0	never	23.55	5.0	73	Female		0	0	former	27.32	7.0
34	Male	5	0	0	<na></na>	15.10	5.8	74	Male		ň	0	former	37.16	9.0
35	Female		0	0	<na></na>	21.76	4.5	75	Female		0	0	never	63.48	8.8
36		26	0	0	never	21.70	6.6	76	Male		ŏ	0		27.32	8.2
37	Male	5	0	0	<na></na>	27.32	6.6	77	Female		ő	0	current	32.27	6.2
38	Female	77	1	1	never	32.02	5.0	78	Female		ŏ	0			7.5
39	Female		0	0	<na></na>	29.30	4.8	79	Female	67	0	0	never	27.32	6.2
39	i eila i e	00	U	0	<na></na>	25.30	4.0	80	Female	80	i	0		27.32	6.8

Cons	ole Ter	minal ×	Background Jobs ×				
R	R 4.3.1 ·	C:/Users/	/ASUS/Downloads/ ≈				
81	Female	NA NA	O	O	never	31./0	6.5
82	Male	80	0	0	never	22.06	9.0
83	Female	47	0	0	never	36.49	7.5
84	Male	53	0	0	current	30.80	6.6
85	Female	61	0	0	not current	39.36	9.0
86	Male	76	0	0	never	31.90	7.5
87	Female	43	0	0	never	26.71	6.5
88	Male	55	0	0	<na></na>	27.32	6.8
89	Male	57	1	1	not current	27.77	6.6
90	Female	43	0	0	never	27.32	6.2
91	Male	63	1	0	ever	35.06	5.8
92	Male	80	0	0	never	23.25	6.1
93	Female	2 70	0	0	current	29.25	8.2
94	Female	42	0	0	never	24.81	9.0
95	Female	80	0	0	former	36.18	6.5
96	Female	52	1	0	never	50.30	6.6
97	Male	71	0	0	never	27.09	8.2
98	Female	43	0	0	never	27.32	6.6
99	Male	71	0	0	never	27.09	8.2
100	Male	80	0	1	former	24.36	7.5
101	Male	59	0	0	current	29.20	8.2
102	Male	29	0	0	current	25.41	6.1
103	Female	68	0	0	<na></na>	40.31	7.5
104	Female	52	0	0	<na></na>	27.32	9.0
105	Male	2 71	0	0	never	26.53	8.8
106	Male	48	1	0	current	36.12	6.8
107	Female	2 79	1	0	former	27.32	6.5
108	Male	37	0	0	never	37.24	7.0
109	Female	73	0	0	never	35.56	5.8
110	Female	59	0	0	former	43.41	6.2
111	Female	80	0	0	never	27.32	6.0
112	Female	e 64	0	0	ever	49.27	8.2
113	Male	43	0	0	never	39.00	8.8
114	Male	43	0	0	never	22.43	7.0
115	Male		0	0	not current	32.19	5.8
116	Male	59	1	0	ever	25.94	9.0
117	Female	43	0	0	never	27.73	8.8
118	Male		0	0	ever	19.46	9.0
	Female		0	0	<na></na>	27.32	8.8
120	Female	43	0	0	<na></na>	27.32	5.8
	blood_	glucos	se_level diabetes				
1			140 0				

R	R 4.3.1 · C:/Users/ASUS/Downlo	oads/ 🖈
120	blood_glucose_level	diahatas
1	140	0
	80	Ö
2 3 4 5 6	158	Ö
4	155	0
5	155	0
6	85	0
7	85	0
8	145	0
9	100	0
10	85	0
11	100	0
12	130	0
13	200	0
14	160	0
15	126	0
16	200	0
17	158	0
18	80	0
19	159	0
20	90	0
21	85	0
22	85	0
23 24	130 140	0
25	126	0
26	80	0
27	158	0
28	155	0
29	200	Ö
30	145	Ö
31	126	Ö
32	100	0
33	85	0
34	85	0
35	130	0
36	200	0
37	130	0
38	159	0
39	159	0
40	160	0

blood_glucose_level	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	40 41 42 43 44 45 46 47 48 49 50 51	3.1 · C:/Users/ASUS/Downloads/ → 160 0 100 0 90 0 159 0 155 0 100 0 155 0 200 0 140 0 158 0 159 0	
80 158 155 155 85 85 145 100 0 88 1 100 2 130 3 200 4 160 5 126	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	41 42 43 44 45 46 47 48 49 50 51	100 0 90 0 159 0 155 0 100 0 155 0 200 0 100 0 140 0 158 0	
158 155 155 156 157 158 158 159 169 100 100 100 100 100 100 100 100 100 10	8 0 6 0 6 0 6 0 6 0 6 0 6 0 0 0 0 0 0 0	42 43 44 45 46 47 48 49 50 51	90 0 159 0 155 0 100 0 155 0 200 0 100 0 140 0	
155 155 155 155 155 155 155 165 16	5 0 5 0 5 0 6 0 6 0 6 0 0 0 0 0 0 0	43 44 45 46 47 48 49 50 51	159 0 155 0 100 0 155 0 200 0 100 0 140 0	
155 85 85 145 100 0 85 1 100 2 130 3 200 4 166 5 126	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	44 45 46 47 48 49 50 51 52	155 0 100 0 155 0 200 0 100 0 140 0 158 0	
85 85 87 88 81 145 9 100 100 85 1.1 100 1.2 130 1.3 200 1.4 166 1.5 126	5 0 5 0 6 0 0 0 0 0 0 0 0 0 0 0	45 46 47 48 49 50 51 52	100 0 155 0 200 0 100 0 140 0 158 0	
85 8 145 9 100 100 85 1.1 100 1.2 130 1.3 200 1.4 166 1.5 126	5 0 5 0 0 0 5 0 0 0 0 0 0 0	46 47 48 49 50 51 52	155 0 200 0 100 0 140 0 158 0	
3 145 0 100 .0 85 .1 100 .2 130 .2 130 .4 166 .5 126	5 0 0 0 5 0 0 0 0 0 0 0	47 48 49 50 51 52	200 0 100 0 140 0 158 0	
100 .0 85 .1 100 .2 130 .3 200 .4 166 .5 126	0 0 5 0 0 0 0 0 0 0	48 49 50 51 52	100 0 140 0 158 0	
.0 85 .1 100 .2 130 .3 200 .4 166 .5 126	0 0 0 0 0 0 0 0	49 50 51 52	140 0 158 0	
.1 100 .2 130 .3 200 .4 160 .5 126	0 0 0 0 0 0	50 51 52	158 0	
.2 130 .3 200 .4 160 .5 126	0 0 0	51 52		
.3 200 .4 160 .5 126	0 0			
.4 160 .5 126	0	_	100 0	
.5 126		53	200 0	
	5 0	54	145 0	
.6 200		55	90 0	
		56	200 0	
158		57	160 0	
.8 80		58	158 0	
.9 159		59	90 0	
20 90		60	145 0	
21 85		61	126 0	
22 85	5 0	62	158 0	
23 130	0	63	130 0	
24 140	0	64	159 0	
25 126	5 0	65	158 0	
26 80	0	66	90 0	
27 158	3 0	67	160 0	
28 155	5 0	68	159 0	
29 200		69	200 1	
30 145	5 0	70	200 1	
31 126	5 0	71	260 1	
32 100		72	160 1	
83 85		73	159 1	
34 85		74	159 1	
35 130		75	155 1	
36 200		76	126 1	
37 130		77	220 1	
38 159		78	300 1	
39 159		79	159 1	
10 160		80	280 1 280 1	
100	·	.81	280 1	
82 15:	5 1			
83 15!	5 1			

82	155	1
83	155	1
84	280	1
85	140	1
86	155	1
87	300	1
88	159	1
89	160	1
90	155	1
91	200	1
92	159	1
93	130	1
94	159	1
95	200	1
96	155	1
97	200	1
98	130	1
99	200	1
100	280	1
101	220	1
102	130	1
103	260	1
104	140	1
105	159	1
106	140	1
107	159	1
108	126	1
109	260	1
110	160	1
111	200	1
112	140	1
113	220	1
114	160	1
115	300	1
116	140	1
117	145	1
118	130	1
119	159	1
120	159	1

Also, there are some missing values based on specific columns and rows. The outputs are:

```
> which(is.na(mydata$gender))
[1] 7 15
> which(is.na(mydata$age))
[1] 4 19 55 81
> which(is.na(mydata$hypertension))
[1] 12 27
> which(is.na(mydata$smoking_history))
[1] 2 7 14 15 18 21 24 29 30 34 35 37 39 40 45 47 48 49 54 55 57 58
[23] 59 61 64 68 76 88 103 104 119 120
> |
```

Data Exploration:

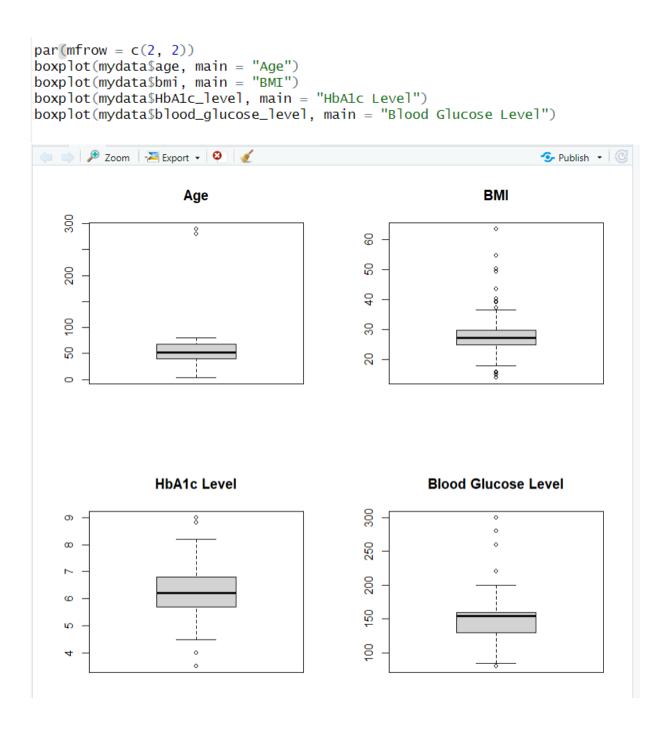
Missing age values, and any negative ages have been converted to their positive counterparts (absolute values):

```
> mydata <- mydata[!is.na(as.numeric(mydata$age)), ]
> unique(mydata$age)
[1] 80 54 28 76 20 79 42 32 53 78 67 15 40 5 69 72 4 30 45 43 50 41
[23] 26 77 66 44 29 60 38 3 57 74 21 59 290 19 56 7 11 34 37 73 36 47
[45] 61 55 63 70 52 71 68 48 64 62 280
> mydata$age[mydata$age < 0] <- abs(mydata$age[mydata$age < 0])</pre>
```

Missing bmi values, and any negative ages have been converted to their positive counterparts (absolute values):

Handling potential outliers using box plots:

This will create a 2x2 grid of boxplots to visualize the distribution and spread of data for four different variables in the mydata dataframe. The output:



From the above boxplots it is noticible that there are few outliers present in the dataset.

Dealing with missing values:

We replaced the main data frame "mydata" each time whenever we performed a specific task on the dataframe so that there would be backup data as mydata. Replacing missing values with the mean value. The output:

```
mydata_mean <- mydata
mean_age<- mean(mydata_mean$age, na.rm = TRUE)
mydata_mean$age[is.na(mydata_mean$age)] <- mean_age
print(mydata_mean)</pre>
```

	gender	age	hypertension	heart_disease	smoking_history	bmi	HbA1c_level	R	R 4.3.1 · C	:/Users/	ASUS/Downloads/ 🖈			
1	Female	80	0	1	never	25.19	6.6	43	⊦ema∣e	bυ	U	U	never 18.03	4.0
2	Female	54	0	0	<na></na>	27.32	6.6	44	Female	38	0	0	never 28.27	6.2
3	Male	28	0	0		27.32	5.7	45	Female	3	0	0	<na> 19.27</na>	6.5
5	Male		1	1	current		4.8	46	Male	57	0	0	never 27.32	6.1
6	Female	20	0	0		27.32	6.6	47	Female	43	0	0	<na> 27.32</na>	4.0
7		79	0	0		23.86	5.7	48	Female	74	0	0	<na> 28.12</na>	5.0
/	<na></na>		0	0				49	Female	21	0	0	<na> 26.10</na>	5.8
8	Male		0	0		33.64	4.8	50	Female	30	0	0	current 27.32	6.5
9	Female		0	0		27.32	5.0	51	Female		0	ō	former 27.32	6.0
10	Female	53	0	0	never	27.32	6.1	52	Female		Ŏ	Ö	not current 30.22	5.7
11	Female	54	0	0	former		6.0	53		59	Õ	1	ever 23.11	6.5
12	Female	78	NA	0	former	36.05	5.0	54		19	Ŏ	ō	<na> 27.32</na>	5.7
13	Female	67	0	0	never	25.69	5.8	56	Male	56	0	Ô	never 26.78	4.8
14	Female	76	0	0	<na></na>	27.32	5.0	57	Male	43	0	0	<na> 23.04</na>	5.7
15	<na></na>	78	0	0	<na></na>	27.32	6.6	58	Male	7	0	0	<na> 25.04 <na> 15.94</na></na>	5.8
16	Male	15	0	0		30.36	6.1	59	Male	3	0	0	<na> 15.80</na>	6.2
17	Female		0	Ō		24.48	5.7	60	Female	30	0	0	never 27.01	6.2
18	Female	42	0	ň		27.32	5.7	61	Male	43	0	0	<na> 27.32</na>	3.5
20	Male	40	0	0	current		6.0	62	Female	76	0	0	never 22.19	6.6
21	Male	5	0	0		18.80	6.2	63	Female	41	0	0	never 27.45	5.7
22	Female	69	0	0		21.24	4.8		Female	11	0	0	<na> 17.98</na>	
			0	1				64		26	0	0	<na> 17.98 never 26.45</na>	6.5 5.7
23	Female	72	0	1	former		6.5	65	Female		0	•		
24	Female	4	0	0		13.99	4.0	66	Male	34	0	0	never 31.16	5.8
25	Male	30	0	0		33.76	6.1	67	Male	80	0	0	former 24.42	4.0
26	Male	40	0	0	former		5.8	68	Female	37	0	0	<na> 30.50</na>	5.7
27	Male	45	NA	0	never	26.47	4.0	69	Female	44	0	0	never 19.31	6.5
28	Male	43	0	0	never	26.08	6.1	70	Male	67	0	1	not current 27.32	6.5
29	Female	53	0	0	<na></na>	31.75	4.0	71	Male	50	1	0	current 27.32	5.7
30	Male	50	0	0	<na></na>	25.15	4.0	72	Male	73	0	0	former 25.91	9.0
31	Female	41	0	0	current	22.01	6.2	73		53	0	0	former 27.32	7.0
32	Female	20	0	0	never	22.19	3.5	74	Male	50	0	0	former 37.16	9.0
33	Female	76	0	0		23.55	5.0	75	Female	67	0	0	never 63.48	8.8
34	Male	5	0	ň		15.10	5.8	76	Male	57	0	0	<na> 27.32</na>	8.2
35	Female	-	0	0		21.76	4.5	77	Female	36	0	0	current 32.27	6.2
36	Female	26	0	0		21.22	6.6	78	Female	60	0	0	never 27.32	7.5
37	Male	5	0	0		27.32	6.6	79	Female	67	0	0	never 27.32	6.2
	Female		0	0				80	Female	80	1	0	never 27.32	6.8
38			1	1		32.02	5.0	82	Male	80	0	0	never 22.06	9.0
39	Female	66	0	0		29.30	4.8	83	Female	47	0	0	never 36.49	7.5
40	Female	67	0	0		27.32	3.5	84		53	0	0	current 30.80	6.6
41	Female	44	0	0		24.93	6.1	85	Female	61	0	0	not current 39.36	9.0
47	Female	79		0	never	19 95	5.0	86	Male	76	0	0	never 31.90	7.5

							Con	sole Terminal	× Background Jobs ×	
8/	Female	43	0	0	never	26./1		R 4.3.1 · C:/Use	rs/ASUS/Downloads/ 🖈	
88	Male	55	Ö	0		27.32	6.8		145	0
89	Male	57	ĭ	ĭ	not current		6.6 ⁹		100	0
90	Female	43	0	0		27.32	6.2		85	0
91	Male	63	ĭ	0		35.06	- 0		100	0
92	Male	80	0	0		23.25	6.1		130	0
93	Female	70	0	0	current				200	0
94		42	0	0		24.81	14		160	0
95	Female	80	0	0	former				126	0
96	Female	52	1	0		50.10	10		200	0
		71							158	0
97	Male		0	0		27.09	8.2		80 90	0
98	Female	43	0	0		27.32	6.6 20			0
99	Male	71	0	0		27.09	8.2 21 7.5 22		85 85	0
100	Male	80	0	1	former				130	0
101	Male	59	0	0	current		8.2 23 6.1 24		140	0
102		29	0	0	current		0.1		140	0
	Female	68	0	0		40.31	7.3		80	0
	Female	52	0	0		27.32	9.0		158	0
105	Male	71	0	0		26.53	8.8		155	0
106		48	1	0	current	36.12	6.8		200	0
107	Female	79	1	0	former	27.32	6.5		145	0
108	Male	37	0	0		37.24	7.0		126	0
109	Female	73	0	0	never	35.56	5.8		100	0
110	Female	59	0	0	former	43.41	6.2		85	0
111	Female	80	0	0	never	27.32	6.0		85	0
112	Female	64	0	0	ever	49.27	8.2 35		130	0
113	Male	43	0	0	never	39.00	8.8 36		200	0
114	Male	43	0	0	never	22.43	7.0		130	0
115	Male	62	0	0	not current		5.8 38		159	0
116		59	1	0		25.94	9.0 39		159	0
	Female	43	0	0		27.73	8.8 40		160	0
118		43	0	0		19.46	9.0 41		100	0
	Female		0	0		27.32	8.8 42		90	0
	Female		0	0		27.32	5.8 43		159	0
120			_level diab		NA.	21.52	3.0		155	0
1	5 Toou_g	,, acost	140	0			45		100	0
2			80	0			46		155	0
3			158	0			47		200	0
5			155	0			48		100	0
			85				49		140	0
6			85 85	0			50		158	٥
			- `							

Console	Terminal × Background Jobs ×		95	200
R 4.	3.1 · C:/Users/ASUS/Downloads/ ↔		96 97	155
51	159	U		200
52	100	0	98	130
53	200	0	99	200
54	145	0	100	280
56	200	0	101	220
57	160	0	102	130
58	158	0	103	260
59	90	0	104	140
60	145	0	105	159
61	126	0	106	140
62	158	0	107	159
63	130	0	108	126
64	159	0	109	260
65	158	0	110	160
66	90	0	111	200
67	160	0	112	140
68	159	0	113	220
69	200	1	114	160
70	200	1	115	300
71	260	1	116	140
72	160	1	117	145
73	159	1	118	130
74	159	1	119	159
75	155	1	120	159
76	126	1	>	
77	220	1		
78	300	1		
79	159	1		
80	280	1		
82	155	1		
83	155	1		
84	280	1		
85	140	1		
86	155	1		
87	300	1		
88	159	1		
89	160	1		
90	155	1		
91	200	1		
92	159	1		
93	130	1		
94	159	1		

Check for missing values and count:

Data Visualization:

ggplot2 library, which was used for creating data visualizations. We replaced mydata data frame with mydata_remove. Then we removed rows with missing values, ensuring that the dataset used for analysis and visualization is complete and doesn't introduce bias or inaccuracies due to missing data. The output:

> ·	library(ggplo	t2)					R	P 4 2 1 . C	/Usors/	ASUS/Downloads/ 🗇				
> r	nydata_re	emove	<-mydata					65	Female		0	0	never	26. 45	5.7
> r	nydata_re	emove	<- mydata_remove	[complete.case	s(mydata_remo	ve),		66	Male		0	0	never		5.8
> 1	nydata_re	emove						67	Male	80	0	0	former		4.0
	gender	age	hypertension hear	t_disease smok	ing_history	bmi	Ula Adia Tarrad	69	Female	44	0	0	never		6.5
1	Female	80	0	1	never 2	5.19		70	Male	67	0	ĭ	not current		6.5
3	Male	28	0	0	never 2	7.32		71	Male	50	1	0	current		5.7
5	Male	76	1	1	current 2	0.14		72	Male	73	0	Õ	former		9.0
6	Female		0	0	never 2			73	Female		0	ŏ	former		7.0
8			Ō	0	never 3			74	Male	50	Ō	ō	former		9.0
9	Female		Ô	0	never 2			75	Female	67	0	0	never	63.48	8.8
10			Ŏ	Ö	never 2			77	Female	36	0	0	current	32.27	6.2
11			ŏ	ŏ	former 5			78	Female	60	0	0	never	27.32	7.5
13			0	0	never 2			79	Female	67	0	0	never	27.32	6.2
16	Male		0	0	never			80	Female	80	1	0	never	27.32	6.8
17	Female		0	0	never 2		5.7	82	Male	80	0	0	never	22.06	9.0
20			0	0	current		6.0	83	Female	47	0	0	never	36.49	7.5
22	Female		0	0	never 2		4.8	84	Male	53	0	0	current	30.80	6.6
23			0	1	former 2		6.5	85	Female	61	0	0	not current	39.36	9.0
			0	_				86	Male	76	0	0	never	31.90	7.5
25	Male		0	0	never			87	Female	43	0	0	never	26.71	6.5
26			0	0	former 2			89	Male	57	1	1	not current		6.6
28	Male		0	0	never 2			90	Female	43	0	0	never		6.2
31	Female		0	0	current 2			91	Male	63	1	0		35.06	5.8
32	Female		0	0	never 2			92	Male	80	0	0	never		6.1
33	Female		0	0	never 2			93	Female	70	0	0	current		8.2
36			0	0	never 2		0.0	94	Female	42	0	0	never		9.0
38	Female		1	1	never		3.0	95	Female	80	0	0	former		6.5
41	Female	44	0	0	never 2	4.93	0.1	96	Female	52	1	0	never		6.6
42	Female	29	0	0	never 1	9.95	3.0	97	Male	71	0	0	never		8.2
43	Female	60	0	0	never 1	8.03		98	Female	43	0	0	never		6.6
44	Female	38	0	0	never 2	8.27		99	Male	71	0	0	never		8.2
46	Male	57	0	0	never 2	7.32		100	Male		0	1	former		7.5
50	Female	30	0	0	current 2	7.32		101	Male	59	0	0	current		8.2
51	Female	59	0	0	former 2	7.32		102	Male	29	0	0	current		6.1
52	Female		Õ	0	not current 3		5 7	105	Male	71	0	0	never		8.8
53	Female		Ŏ	1	ever 2		6.5	106	Male	48	1	0	current		6.8
56			Õ	0	never 2		4 0		Female	79	1	0	former		6.5
60	Female		0	0	never 2		6.3	108	Male	37	0	0	never		7.0
62			0	0	never 2				Female	73	0	0	never		5.8
63	Female		0	0	never 2				Female	59	0	0	former		6.2
65			0	0	never 2				Female	80	0	0	never	40.37	6.0
- 63	- PINIA I D	211			ADODI:	41	1 /								

						onsole	Terminal × Ba	ackground Jobs ×	
Console Terminal × Ba	//Pounloads/						1 · C:/Users/ASU	-	
.12 Female 64	0	0	ever 49.27	8.2	6	-		158	ŏ
.13 Male 43	0	0	never 39.00	8.8	6			130	0
.14 Male 43	0	0	never 22.43	7.0	6			158	0
15 Male 62	0	0	not current 32.19	5.8	6			90	0
16 Male 59	1	0	ever 25.94	9.0	6			160	0
L7 Female 43	0	0	never 27.73	8.8	6			200	1
.8 Male 43	0	0	ever 19.46	9.0	7			200	1
blood_glucose_l	140 0				7			260	
	158 0								1
	155 0				7			160	1
	85 0				7			159	1
	145 0				7			159	1
	100 0				7			155	1
)	85 0				7			220	1
L	100 0				7			300	1
3	200 0				7			159	1
	200 0				8	0		280	1
	158 0				8	2		155	1
)	90 0				8	3		155	1
	85 0 130 0				8			280	1
	126 0				8			140	1
	80 0				8			155	1
	155 0				8			300	1
	126 0				8			160	1
!	100 0				9			155	1
	85 0				9			200	1
i	200 0				9			159	1
	159 0				9			130	
	100 0								1
	90 0				9			159	1
	159 0				9			200	1
	155 0 155 0				9			155	1
i I	158 0				9			200	1
	159 0				9			130	1
	100 0				9			200	1
	200 0					.00		280	1
i	200 0					.01		220	1
)	145 0				1	.02		130	1
					1	.05		159	1
					1	.06		140	1
					1	.07		159	1
						.08		126	1
						.09		260	1

```
    110
    160

    111
    200

    112
    140

    113
    220

    114
    160

    115
    300

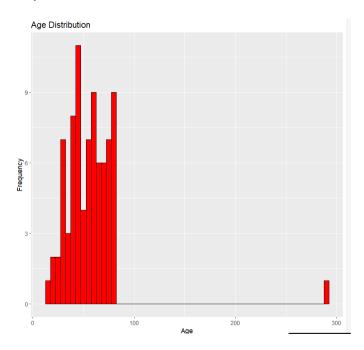
    116
    140

    117
    145

    118
    130
```

To indentify pattern trend in the dataset .The output:

```
> ggp!ot(mydata_remove, aes(x = age)) +
+ geom_histogram(binwidth = 5, fill = "red", color = "black") +
+ labs(x = "Age", y = "Frequency", title = "Age Distribution")
> |
```



Checking for missing values and count. The output:

```
missing_count <- colSums(is.na(mydata))
missing_count

gender age hypertension heart_disease
2 0 2 0
smoking_history bmi HbAlc_level blood_glucose_level
31 0 0 0
diabetes</pre>
```

Replacing missing values in the column with the mode. The output:

R 4.3.1 · C:/Users/ASUS/Downloa	ads/ 🖈					sole Term		round Jobs ×			
> mydata_mode <- mydata				TDUE \ \ [1]	R	R 4.3.1 · C	:/Users/ASUS/Do	wnloads/ 🔑		Carrent LL.OL	·
<pre>> mode_gender <- names(so > mydata_mode\$gender[is.r</pre>				TRUE))[1]	32	Female		0	0	never 22.19	3.5
> mode_age <- names(sort)				[1]	33 34	Female Male	76 5	0	0	never 23.55 never 15.10	5.0 5.8
<pre>> mydata_mode\$age[is.na(r</pre>	mydata_mode\$age)]	<- mode_age			35	Female	15	0	0	never 21.76	4.5
> mode_hypertension <- na					36	Female	26	0	0	never 21.22	6.6
<pre>> mydata_mode\$hypertensic > mode_hsmoking <- names</pre>					37 38	Male Female	5 77	0	0 1	never 27.32	6.6 5.0
> mydata_mode\$smoking_his						Female	66	1 0	0	never 32.02 never 29.30	4.8
> mydata_mode\$smoking_his	story[is.na(mydat	a_mode\$smoking_h	istory)] <-	mode_hsmokin	40	Female	67	0	0	never 27.32	3.5
> mydata_mode gender age hypertensi	ion hoart disease	emoking history	bmi ⊎b∧1	c lovel	41 42	Female Female	44 29	0	0	never 24.93 never 19.95	6.1 5.0
1 Female 80	0 1		25.19	6.6	43	Female	60	0	0	never 18.03	4.0
2 Female 54	0 (never never	27.32	6.6	44	Female	38	0	0	never 28.27	6.2
3 Male 28	0 (5.7	45 46	Female Male	3 57	0	0	never 19.27 never 27.32	6.5 6.1
5 Male 76 6 Female 20	1 1			4.8 6.6	47	Female	43	0	0	never 27.32	4.0
7 Female 79	0 (never	23.86	5.7	48	Female	74	0	0	never 28.12	5.0
8 Male 42	0 (4.8	49 50	Female Female	21 30	0	0	never 26.10 current 27.32	5.8 6.5
9 Female 32 10 Female 53	0 0			5.0 6.1	51	Female	59	0	o	former 27.32	6.0
11 Female 54	ŏ			6.0	52	Female		0	0	not current 30.22	5.7
12 Female 78	0 (5.0	53 54	Female Female	59 19	0	1	ever 23.11 never 27.32	6.5 5.7
13 Female 67 14 Female 76	0 0			5.8 5.0	56	Male	56	ŏ	ŏ	never 26.78	4.8
15 Female 78	0 0			6.6	57	Male	43	0	0	never 23.04	5.7
16 Male 15	0 (6.1	58 59	Male Male	7 3	0	0	never 15.94 never 15.80	5.8 6.2
17 Female 42	0 (5.7	60	Female	30	Ö	ő	never 27.01	6.2
18 Female 42 20 Male 40	0 (5.7 6.0	61	Male	43	0	0	never 27.32	3.5
21 Male 5	0 (never never	18.80	6.2	62 63	Female Female	76 41	0	0	never 22.19 never 27.45	6.6 5.7
22 Female 69	0 (4.8	64	Female	11	0	ō	never 17.98	6.5
23 Female 72 24 Female 4	0 1			6.5 4.0	65 66	Female Male	26 34	0	0	never 26.45 never 31.16	5.7
25 Male 30	0 (never		6.1	67	Male	80	0	0	former 24.42	5.8 4.0
26 Male 40	0 (5.8	68	Female	37	0	0	never 30.50	5.7
27 Male 45 28 Male 43	0 0		26.4/	4.0 6.1	69 - 70	Female Male	44 67	0	0 1	never 19.31 not current 27.32	6.5 6.5
29 Female 53	ŏ č		31.75	4.0	1]71	Male	50	1	0	current 27.32	5.7
30 Male 50 31 Female 41	0 0			4.0 6.2	72 73	Male Female	73 53	0	0	former 25.91 former 27.32	9.0 7.0
74 Male 50 75 Female 67 76 Male 57 77 Female 36 78 Female 36 78 Female 67 79 Female 67 80 Female 80 82 Male 80 83 Female 47 84 Male 53 85 Female 61 86 Male 76 87 Female 43 88 Male 55 89 Male 55 89 Male 57 90 Female 43 91 Male 80 93 Female 43 91 Female 43 92 Male 80 93 Female 43 91 Male 80 93 Female 43 91 Male 80 93 Female 43 91 Male 80 93 Female 42 95 Female 80 96 Female 52 97 Male 71 98 Female 43 99 Male 71 98 Female 68 104 Female 79 108 Male 71 106 Male 80 107 Female 79 108 Male 71	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		37.16 63.48 27.32 32.27 27.32 27.32 27.32 27.32 22.06 36.49 30.80 39.36 31.90 26.71 27.32 27.77 27.32 35.06 23.25 29.25 24.81 36.18 36.18 27.09 27.79 27.32 27.77 27.32 37.24 36.18 36.18 36.18 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 27.39 24.36 29.20 25.41 40.31 27.32 26.53 36.12 27.32 27.32 37.24 35.56	9.0 8.8 8.2 6.2 7.5 6.2 6.8 9.0 7.5 6.8 6.6.2 5.8 8.2 9.0 6.5 6.6.8 6.2 5.8 8.2 9.0 6.5 6.8 6.2 6.8 6.2 6.8 6.9 6.9 6.9 6.9 6.9 6.9 6.9 6.9	118	7 Female 3 Male 9 Female 9 Female	43 43 280 43 280 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0	ever 25.94 never 27.73 ever 19.46 never 27.32 never 27.32	9.0 8.8 9.0 8.8 5.8
111 Female 80 112 Female 64	0) ever	27.32 49.27	6.0 8.2	33			85 0			
113 Male 43 114 Male 43			39.00 22.43	8.8 7.0	34 35			85 0 130 0			
115 Male 62	0	not current		5.8	36			200 0			

Console	Terminal ×	Background Jobs ×	
R R 4.	3.1 · C:/Users/A	ASUS/Downloads/ 🖈	
38		T2A	U
39		159	0
40		160	0
41		100	0
42		90	0
43		159	0
44		155	0
45		100	0
46		155	0
47		200	0
48		100	0
49		140	0
50		158	0
51		159	0
52		100	0
53		200	0
54		145	0
56		200	0
57		160	0
58		158	0
59		90	0
60		145	0
61		126	0
62		158	0
63		130	0
64		159	0
65		158	0
66		90	0
67		160	0
68		159	0
69		200	1
70		200	1
71		260	1
72		160	1
73		159	1
74		159	1
75		155	1
76		126	1
77		220	1
78		300	1
79		159	1
80		280	1

82	155	1
83	155	1
84	280	1
85	140	1
86	155	1
87	300	1
88	159	1
89	160	1
90	155	1
91	200	1
92	159	1
93	130	1
94	159	1
95	200	1
96	155	1
97	200	1
98	130	1
99	200	1
100	280	1
101	220	1
102	130	1
103	260	1
104	140	1
105	159	1
106	140	1
107	159	1
108	126	1
109	260	1
110	160	1
111	200	1
112	140	1
113	220	1
114	160	1
115	300	1
116	140	1
117	145	1
118	130	1
119	159	1
120	159	1

			<- mydata				R	D 4 2 1 C		US/Downloads/ 🖈			
		data_	_median\$age)				39	remale		U U	U	<na> 29.30</na>	4.8
	52.5						40			0	Õ	<na> 27.32</na>	3.5
> m	ydata_me						41	Female	44	o o	Ö	never 24.93	6.1
	gender	age h	hypertension heart_d	disease smok	ing_history br	ni HbA1c_level	42	Female	29	0	ŏ	never 19.95	5.0
1	Female	80	0	1	never 25.1	19 6.6	43	Female	60	o o	Ö	never 18.03	4.0
2	Female	54	0	0	<na> 27.</na>	32 6.6	44	Female		0	ŏ	never 28.27	6.2
3	Male	28	0	0	never 27.	32 5.7	45	Female	3	Ō	ō	<na> 19.27</na>	6.5
5	Male	76	1	1	current 20.1	4.8	46	Male	57	Ō	ō	never 27.32	6.1
6	Female	20	0	0	never 27.		47	Female		Ō	ō	<na> 27.32</na>	4.0
7	<na></na>	79	0	0	<na> 23.8</na>		48	Female	74	Ō	ō	<na> 28.12</na>	5.0
8	Male		0	Ö	never 33.6		49	Female	21	0	0	<na> 26.10</na>	5.8
9	Female		0	0	never 27.		50	Female	30	0	0	current 27.32	6.5
10	Female		Ö	ŏ	never 27.		51	Female	59	0	0	former 27.32	6.0
11	Female		0	Ö	former 54.7		52	Female	290	0	0	not current 30.22	5.7
		78	NA	ő	former 36.0		53	Female	59	0	1	ever 23.11	6.5
	Female		0	0	never 25.6		54	Female	19	0	0	<na> 27.32</na>	5.7
14	Female		0	0	<na> 27.3</na>		56	Male	56	0	0	never 26.78	4.8
15	<na></na>		0	0	<na> 27</na>		57	Male	43	0	0	<na> 23.04</na>	5.7
	Male		0	0	never 30.		58	Male	7	0	0	<na> 15.94</na>	5.8
16	Female		0	0	never 24.4		59	Male	3	0	0	<na> 15.80</na>	6.2
17			0	0	<na> 27.3</na>		60	Female	30	0	0	never 27.01	6.2
18	Female						61	Male	43	0	0	<na> 27.32</na>	3.5
20	Male	40	0	0	current 36.		62	Female	76	0	0	never 22.19	6.6
21	Male	5	0	0	<na> 18.8</na>		63	Female	41	0	0	never 27.45	5.7
22		69	0	0	never 21.2		64	Female	11	0	0	<na> 17.98</na>	6.5
23		72	0	1	former 27.9		65	Female	26	0	0	never 26.45	5.7
24	Female	4	0	0	<na> 13.9</na>		66	Male	34	0	0	never 31.16	5.8
25	Male	30	0	0	never 33.7		67	Male	80	0	0	former 24.42	4.0
26	Male		0	0	former 27.8		68	Female	37	0	0	<na> 30.50</na>	5.7
27	Male		NA	0	never 26.4		69	Female	44	0	0	never 19.31	6.5
28	Male	43	0	0	never 26.0	08 6.1	70	Male	67	0	1	not current 27.32	6.5
29	Female	53	0	0	<na> 31.7</na>	75 4.0	71	Male	50	1	0	current 27.32	5.7
30	Male	50	0	0	<na> 25.1</na>	15 4.0	72	Male	73	0	0	former 25.91	9.0
31	Female	41	0	0	current 22.0	01 6.2	73	Female	53	0	0	former 27.32	7.0
32	Female	20	0	0	never 22.1	19 3.5	74	Male	50	0	0	former 37.16	9.0
33	Female	76	0	0	never 23.5	55 5.0	75	Female	67	0	0	never 63.48	8.8
34	Male	5	0	0	<na> 15.1</na>		76 77	Male	57	0	0	<na> 27.32</na>	8.2
	Female		Ŏ	Ö	<na> 21.7</na>			Female	36	U	_	current 32.27	6.2
36		26	0	0	never 21.2		78	Female	60	0	0	never 27.32 never 27.32	7.5 6.2
37	Male	5	Õ	ő	<na> 27.</na>		79 80	Female Female	67 80	1	0	never 27.32 never 27.32	6.8
	Female		i	ĭ	never 32 (82	Male	80	0	0	never 27.32	9.0

83 Female 47	0	0	never 36.49	7.5	R 4.3.1 ⋅ 0	:/Users/ASUS/Downloads/ 🗇		
84 Male 53	0	0	current 30.80	6.6	3	158	0	
85 Female 61	0	Ō	not current 39.36	9.0	5	155	0	
86 Male 76	0	ō	never 31.90	7.5	6	85	0	
87 Female 43	Ö	ŏ	never 26.71	6.5	7	85	0	
88 Male 55	0	ŏ	<na> 27.32</na>	6.8	8	145	0	
89 Male 57	ĭ	1	not current 27.77	6.6	9	100	0	
90 Female 43	0	0	never 27.32	6.2	10	85	0	
91 Male 63	1	0	ever 35.06	5.8	11	100	0	
92 Male 80	0	0	never 23.25	6.1	12	130	0	
93 Female 70	0	0	current 29.25	8.2	13	200	0	
	0	0	never 24.81		14	160	0	
	0	0		9.0	15	126	0	
95 Female 80	•		former 36.18	6.5	16	200	0	
96 Female 52	1	0	never 50.30	6.6	17	158	0	
97 Male 71	0	0	never 27.09	8.2	18	80	0	
98 Female 43	0	0	never 27.32	6.6	20	90	0	
99 Male 71	0	0	never 27.09	8.2	21	85	0	
100 Male 80	0	1	former 24.36	7.5	22	85	0	
101 Male 59	0	0	current 29.20	8.2	23	130	0	
102 Male 29	0	0	current 25.41	6.1	24	140	0	
103 Female 68	0	0	<na> 40.31</na>	7.5	25	126	0	
104 Female 52	0	0	<na> 27.32</na>	9.0	26	80	0	
105 Male 71	0	0	never 26.53	8.8	27	158	0	
106 Male 48	1	0	current 36.12	6.8	28	155	0	
107 Female 79	1	0	former 27.32	6.5	29	200	0	
108 Male 37	0	Ō	never 37.24	7.0	30	145	0	
109 Female 73	Ö	Ö	never 35.56	5.8	31	126	0	
110 Female 59	0	Ō	former 43.41	6.2	32	100	0	
111 Female 80	Ö	Õ	never 27.32	6.0	33	85	0	
112 Female 64	0	0	ever 49.27	8.2	34	85	0	
113 Male 43	0	0	never 39.00	8.8	35	130	0	
113 Male 43 114 Male 43	0	0	never 22.43	7.0	36	200	0	
114 Male 45 115 Male 62	0	0	not current 32.19	5.8	37	130	0	
115 Male 62 116 Male 59	1	0	ever 25.94		38	159	0	
	1	_		9.0	39	159	0	
117 Female 43	0	0	never 27.73	8.8	40	160	0	
118 Male 43	0	0	ever 19.46	9.0	41	100	0	
119 Female 280	0	0	<na> 27.32</na>	8.8	42	90	0	
120 Female 43	. 0	0	<na> 27.32</na>	5.8	43	159	0	
blood_glucose_					44	155	0	
1	140 0				45 46	100	0	
2	٥٥ ٥				.16	155		

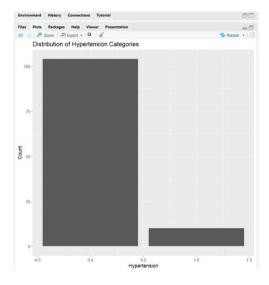
47	200	0
48	100	0
49	140	0
50	158	0
51	159	0
52	100	ō
53	200	ō
54	145	ő
56	200	ő
57	160	Ö
58	158	0
59	90	0
60	145	0
61	126	0
62	158	0
63	130	0
64	159	0
65	158	0
66	90	0
67	160	0
68	159	0
69	200	1
70	200	1
71	260	1
72	160	1
73	159	1
74	159	1
75	155	1
76	126	1
77	220	1
78	300	1
79	159	1
80	280	1
82	155	1
83	155	1
84	280	1
85	140	1
86	155	1
87	300	ī
88	159	ī
00	100	-

90	155	1
91	200	1
92	159	1
93	130	1
94	159	1
95	200	1
96	155	1
97	200	1
98	130	1
99	200	1
100	280	1
101	220	1
102	130	1
103	260	1
104	140	1
105	159	1
106	140	1
107	159	1
108	126	1
109	260	1
110	160	1
111	200	1
112	140	1
113	220	1
114	160	1
115	300	1
116	140	1
117	145	1
118	130	1
119	159	1
120	159	1

Data cleaning by counting the total missing column. The output:

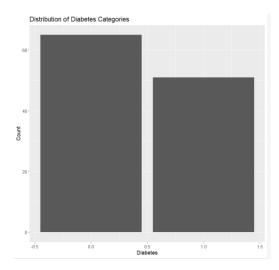
A visual representation of the counts of different hypertension categories, that will help to understand the distribution of this variable in mydata. The output:

```
C.\USEIS\ASUS\APPDACA\LOCAI\TEMP\KCMPYNELNK\UOWNITUAUEU_PACKAYES
> library(ggplot2)
Warning message:
package 'ggplot2' was built under R version 4.3.2
> ggplot(mydata, aes(x = hypertension)) +
+ geom_bar() +
+ labs(x = "Hypertension", y = "Count", title = "Distribution of Hypertension Categories")
```



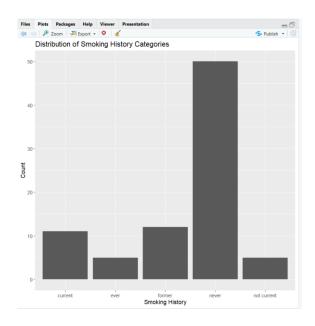
A visual representation of the counts of different diabetes categories, that will help to understand the distribution of this variable in my data. The output:

```
> ggplot(mydata, aes(x = dlabetes)) +
+    geom_bar() +
+    labs(x = "Diabetes", y = "Count", title = "Distribution of Diabetes Categories")
> |
```

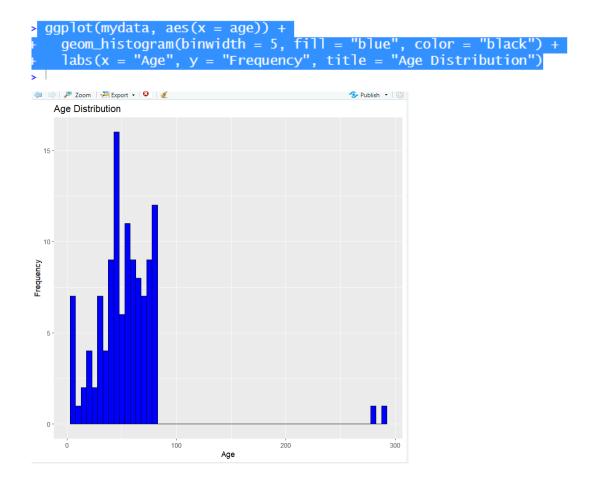


A visual representation of the counts of different heart disease categories. The output:

A visual representation of the counts of different smoking histories categories. The output:



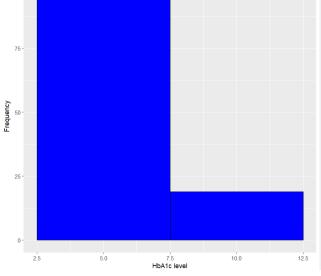
Ages are distributed across different ranges. The output:



The BMIs are distributed across different ranges. The output:

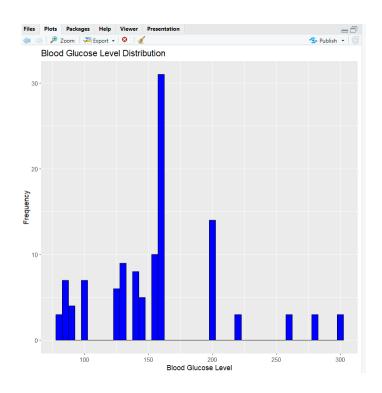
```
ggplot(mydata, aes(x = bmi)) +
  geom_histogram(binwidth = 5, fill = "blue", color = "black") +
    labs(x = "BMI", y = "Frequency", title = "BMI Distribution")
   BMI Distribution
                                                                  nges. The output:
                                                                lor = "black") +
                                                                   = "HbA1x Level Distribution")
Files Plots Packages Help Viewer Presentation

(□ □ ▶ 2 Zoom - Export • □ ✓
   HbA1x Level Distribution
 75
```



The Blood glucose levels are distributed across different ranges. The output:

```
ggplot(mydata, aes(x = blood_glucose_level)) +
  geom_histogram(binwidth = 5, fill = "blue", color = "black") +
  labs(x = "Blood Glucose Level", y = "Frequency", title = "Blood Glucose Level Distribution")
```



After missing values handling:

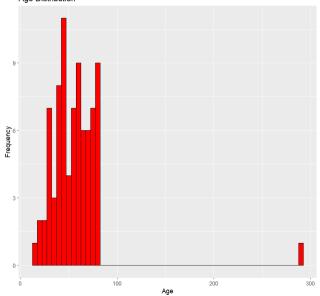
Numerical values of ages are distributed across different age ranges in 'mydata_remove' dataset. The output:

```
ggplot(mydata_remove, aes(x = age)) +
geom_histogram(binwidth = 5, fill = "red", color = "black") +
labs(x = "Age", y = "Frequency", title = "Age Distribution")

Files Plots Packages Help Viewer Presentation

Age Distribution

Age Distribution
```



Numerical values of BMIs are distributed across different bmi ranges in my 'mydata_remove' dataset. The output:

```
ggplot(mydata_remove, aes(x = bmi)) +
  geom_histogram(binwidth = 5, fill = "red", color = "black") +
  labs(x = "BMI", y = "Frequency", title = "BMI Distribution")

BMI Distribution

**Publish ** **

**Publish ** **

**BMI Distribution**

**BMI Distribution*
```

How HBA1cs are distributed between two different ranges in my 'mydata_remove' dataset. The output:

This is how an effective comparison can be identified by performing data exploration, outliers' detection, missing values detection and handling the missing values by implementing mean, mode, median and row wise removing techniques.