

<b>Write-up</b>	<b>Correctness of Program</b>	<b>Documentation of Program</b>	<b>Viva</b>	<b>Timely Completion</b>	<b>Total</b>	<b>Dated Sign of Subject Teacher</b>
<b>4</b>	<b>4</b>	<b>4</b>	<b>4</b>	<b>4</b>	<b>20</b>	

---

## **Group A**

### **Assignment No: 1**

---

**Title of the Assignment: Data Wrangling, I**

Perform the following operations using Python on any open source dataset

(e.g., data.csv) Import all the required Python Libraries.

1. Locate open source data from the web (e.g. <https://www.kaggle.com>).
2. Provide a clear description of the data and its source (i.e., URL of the web site).
3. Load the Dataset into the pandas data frame.
4. Data Preprocessing: check for missing values in the data using pandas `isnull()`, `describe()` function to get some initial statistics. Provide variable descriptions. Types of variables etc. Check the dimensions of the data frame.
5. Data Formatting and Data Normalization: Summarize the types of variables by checking the data types (i.e., character, numeric, integer, factor, and logical) of the variables in the data set. If variables are not in the correct data type, apply proper type conversions.
6. Turn categorical variables into quantitative variables in Python.

---

**Output:**

My Quick Converter | kaggle - Saferbrowser Yahoo India | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 0 seconds To save and continue working Sign In or Register

notebook2da9e7a91a

File Edit View Run Help

+ X D Run All Code Draft Session (2m)

[3]: 

```
import numpy as np # linear algebra
import pandas as pd
```

[4]: 

```
df=pd.read_csv("../input/student-performance-data-set/student-por.csv")
print("read the csv file")
```

read the csv file

[5]: 

```
df.head()
```

[5]:

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	...	4	3	4	1	1	3	4	0	11	11
1	GP	F	17	U	GT3	T	1	1	at_home	other	...	5	3	3	1	1	3	2	9	11	11
2	GP	F	15	U	LE3	T	1	1	at_home	other	...	4	3	2	2	3	3	6	12	13	12
3	GP	F	15	U	GT3	T	4	2	health	services	...	3	2	2	1	1	5	0	14	14	14
4	GP	F	16	U	GT3	T	3	3	other	other	...	4	3	2	1	2	5	0	11	13	13

5 rows × 33 columns

Console 2:15 PM 14/03/2022

Type here to search

My Quick Converter | kaggle - Saferbrowser Yahoo India | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 0 seconds To save and continue working Sign In or Register

notebook2da9e7a91a

File Edit View Run Help

+ X D Run All Code Draft Session (4m)

[6]: 

```
df.tail()
```

[6]:

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3
644	MS	F	19	R	GT3	T	2	3	services	other	...	5	4	2	1	2	5	4	10	11	10
645	MS	F	18	U	LE3	T	3	1	teacher	services	...	4	3	4	1	1	1	4	15	15	16
646	MS	F	18	U	GT3	T	1	1	other	other	...	1	1	1	1	1	5	6	11	12	9
647	MS	M	17	U	LE3	T	3	1	services	services	...	2	4	5	3	4	2	6	10	10	10
648	MS	M	18	R	LE3	T	3	2	services	other	...	4	4	1	3	4	5	4	10	11	11

5 rows × 33 columns

Console 2:17 PM 14/03/2022

Type here to search

My Quick Converter | kaggle - Saferbrowser Yahoo India | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 0 seconds To save and continue working Sign In or Register

notebook2da9e7a91a

File Edit View Run Help

+ X D Run All Code Draft Session (4m)

[7]: 

```
df.describe()
```

[7]:

	age	Medu	Fedu	traveltime	studytime	failures	famrel	freetime	goout	Dalc	Walc	health	absences	G1
count	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000
mean	16.744222	2.514638	2.306626	1.568567	1.930663	0.221880	3.930663	3.180277	3.184900	1.502311	2.280431	3.536210	3.659476	11.399076
std	1.218138	1.124552	1.099931	0.748660	0.829510	0.593235	0.955717	1.051093	1.175766	0.924834	1.284380	1.446759	4.640759	2.745265

Console 2:17 PM 14/03/2022

Type here to search

Kaggle Notebook Session

Logged out session ends in 0 seconds

notebook2da9e7a91a Draft saved

File Edit View Run Help

Code

648 MS M 18 R LE3 T 3 2 services other ... 4 4 1 3 4 5 4 10 11 11

5 rows x 33 columns

[7]: df.describe()

```
age      Medu     Fedu   traveltime  studytime  failures  famrel  freetime  gout    Dalc    Walc  health  absences  G1
count  649.000000 649.000000 649.000000 649.000000 649.000000 649.000000 649.000000 649.000000 649.000000 649.000000 649.000000
mean   16.744222 2.514638 2.036626 1.568567 1.930663 0.221880 3.920663 3.184900 1.502311 2.280431 3.536210 3.659476 11.399076
std    1.218138 1.134552 1.099931 0.748660 0.829510 0.593235 0.955717 1.051093 1.175766 0.924834 1.284380 1.446259 4.640759 2.745265
min    15.000000 0.000000 0.000000 1.000000 1.000000 0.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 0.000000 0.000000
25%   16.000000 2.000000 1.000000 1.000000 1.000000 4.000000 3.000000 2.000000 1.000000 1.000000 2.000000 4.000000 10.000000
50%   17.000000 2.000000 2.000000 1.000000 2.000000 0.000000 4.000000 3.000000 1.000000 2.000000 3.000000 5.000000 6.000000 13.000000
75%   18.000000 4.000000 3.000000 2.000000 2.000000 5.000000 4.000000 4.000000 2.000000 3.000000 5.000000 5.000000 32.000000 19.000000
max   22.000000 4.000000 4.000000 4.000000 4.000000 3.000000 5.000000 5.000000 5.000000 5.000000 5.000000 5.000000 32.000000 19.000000
```

Console

Type here to search

My Quick Converter | kaggle - Saferbrowser Yahoo Inc | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 0 seconds

notebook2da9e7a91a Draft Session (6m)

File Edit View Run All Code

Input

- student-performance-data-set
  - student-por.csv
- autosdataset
  - autos.csv

Output (6MB / 1.9GB)

/kaggle/working

Settings

Code Help

FIND CODE HELP

Find Code Help

Search for examples of how to do things

[8]: df=df.rename(columns={'school':'college'})  
df.info()

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 649 entries, 0 to 648  
Data columns (total: 33 columns):  
 #   Column          Non-Null Count  Dtype     
---    
 0   college        649 non-null   object    
 1   sex            649 non-null   object    
 2   age            649 non-null   int64     
 3   address        649 non-null   object    
 4   famsize        649 non-null   object    
 5   status          649 non-null   object    
 6   Medu           649 non-null   int64     
 7   Fedu           649 non-null   int64     
 8   Mjob            649 non-null   object    
 9   Fjob            649 non-null   object    
 10  reason          649 non-null   object    
 11  guardian        649 non-null   object    
 12  traveltime      649 non-null   int64     
 13  studytime       649 non-null   int64     
 14  failures        649 non-null   int64     
 15  schoolsup       649 non-null   object    
 16  famsup          649 non-null   object    
 17  paid             649 non-null   object    
 18  activities       649 non-null   object    
 19  nursery          649 non-null   object    
 20  higher           649 non-null   object    
 21  internet         649 non-null   object    
 22  romantic         649 non-null   object    
 23  famrel           649 non-null   int64     
 24  G1              649 non-null   int64     
 25  G2              649 non-null   int64     
 26  G3              649 non-null   int64     
 27  G4              649 non-null   int64     
 28  G5              649 non-null   int64     
 29  G6              649 non-null   int64     
 30  G7              649 non-null   int64     
 31  G8              649 non-null   int64     
 32  G9              649 non-null   int64     
 33  G10             649 non-null   int64     
 34  G11             649 non-null   int64     
 35  G12             649 non-null   int64     
 36  G13             649 non-null   int64     
 37  G14             649 non-null   int64     
 38  G15             649 non-null   int64     
 39  G16             649 non-null   int64     
 40  G17             649 non-null   int64     
 41  G18             649 non-null   int64     
 42  G19             649 non-null   int64     
 43  G20             649 non-null   int64     
 44  G21             649 non-null   int64     
 45  G22             649 non-null   int64     
 46  G23             649 non-null   int64     
 47  G24             649 non-null   int64     
 48  G25             649 non-null   int64     
 49  G26             649 non-null   int64     
 50  G27             649 non-null   int64     
 51  G28             649 non-null   int64     
 52  G29             649 non-null   int64     
 53  G30             649 non-null   int64     
 54  G31             649 non-null   int64     
 55  G32             649 non-null   int64     
 56  G33             649 non-null   int64     
 57  G34             649 non-null   int64     
 58  G35             649 non-null   int64     
 59  G36             649 non-null   int64     
 60  G37             649 non-null   int64     
 61  G38             649 non-null   int64     
 62  G39             649 non-null   int64     
 63  G40             649 non-null   int64     
 64  G41             649 non-null   int64     
 65  G42             649 non-null   int64     
 66  G43             649 non-null   int64     
 67  G44             649 non-null   int64     
 68  G45             649 non-null   int64     
 69  G46             649 non-null   int64     
 70  G47             649 non-null   int64     
 71  G48             649 non-null   int64     
 72  G49             649 non-null   int64     
 73  G50             649 non-null   int64     
 74  G51             649 non-null   int64     
 75  G52             649 non-null   int64     
 76  G53             649 non-null   int64     
 77  G54             649 non-null   int64     
 78  G55             649 non-null   int64     
 79  G56             649 non-null   int64     
 80  G57             649 non-null   int64     
 81  G58             649 non-null   int64     
 82  G59             649 non-null   int64     
 83  G60             649 non-null   int64     
 84  G61             649 non-null   int64     
 85  G62             649 non-null   int64     
 86  G63             649 non-null   int64     
 87  G64             649 non-null   int64     
 88  G65             649 non-null   int64     
 89  G66             649 non-null   int64     
 90  G67             649 non-null   int64     
 91  G68             649 non-null   int64     
 92  G69             649 non-null   int64     
 93  G70             649 non-null   int64     
 94  G71             649 non-null   int64     
 95  G72             649 non-null   int64     
 96  G73             649 non-null   int64     
 97  G74             649 non-null   int64     
 98  G75             649 non-null   int64     
 99  G76             649 non-null   int64     
 100  G77             649 non-null   int64     
 101  G78             649 non-null   int64     
 102  G79             649 non-null   int64     
 103  G80             649 non-null   int64     
 104  G81             649 non-null   int64     
 105  G82             649 non-null   int64     
 106  G83             649 non-null   int64     
 107  G84             649 non-null   int64     
 108  G85             649 non-null   int64     
 109  G86             649 non-null   int64     
 110  G87             649 non-null   int64     
 111  G88             649 non-null   int64     
 112  G89             649 non-null   int64     
 113  G90             649 non-null   int64     
 114  G91             649 non-null   int64     
 115  G92             649 non-null   int64     
 116  G93             649 non-null   int64     
 117  G94             649 non-null   int64     
 118  G95             649 non-null   int64     
 119  G96             649 non-null   int64     
 120  G97             649 non-null   int64     
 121  G98             649 non-null   int64     
 122  G99             649 non-null   int64     
 123  G100            649 non-null   int64     
 124  G101            649 non-null   int64     
 125  G102            649 non-null   int64     
 126  G103            649 non-null   int64     
 127  G104            649 non-null   int64     
 128  G105            649 non-null   int64     
 129  G106            649 non-null   int64     
 130  G107            649 non-null   int64     
 131  G108            649 non-null   int64     
 132  G109            649 non-null   int64     
 133  G110            649 non-null   int64     
 134  G111            649 non-null   int64     
 135  G112            649 non-null   int64     
 136  G113            649 non-null   int64     
 137  G114            649 non-null   int64     
 138  G115            649 non-null   int64     
 139  G116            649 non-null   int64     
 140  G117            649 non-null   int64     
 141  G118            649 non-null   int64     
 142  G119            649 non-null   int64     
 143  G120            649 non-null   int64     
 144  G121            649 non-null   int64     
 145  G122            649 non-null   int64     
 146  G123            649 non-null   int64     
 147  G124            649 non-null   int64     
 148  G125            649 non-null   int64     
 149  G126            649 non-null   int64     
 150  G127            649 non-null   int64     
 151  G128            649 non-null   int64     
 152  G129            649 non-null   int64     
 153  G130            649 non-null   int64     
 154  G131            649 non-null   int64     
 155  G132            649 non-null   int64     
 156  G133            649 non-null   int64     
 157  G134            649 non-null   int64     
 158  G135            649 non-null   int64     
 159  G136            649 non-null   int64     
 160  G137            649 non-null   int64     
 161  G138            649 non-null   int64     
 162  G139            649 non-null   int64     
 163  G140            649 non-null   int64     
 164  G141            649 non-null   int64     
 165  G142            649 non-null   int64     
 166  G143            649 non-null   int64     
 167  G144            649 non-null   int64     
 168  G145            649 non-null   int64     
 169  G146            649 non-null   int64     
 170  G147            649 non-null   int64     
 171  G148            649 non-null   int64     
 172  G149            649 non-null   int64     
 173  G150            649 non-null   int64     
 174  G151            649 non-null   int64     
 175  G152            649 non-null   int64     
 176  G153            649 non-null   int64     
 177  G154            649 non-null   int64     
 178  G155            649 non-null   int64     
 179  G156            649 non-null   int64     
 180  G157            649 non-null   int64     
 181  G158            649 non-null   int64     
 182  G159            649 non-null   int64     
 183  G160            649 non-null   int64     
 184  G161            649 non-null   int64     
 185  G162            649 non-null   int64     
 186  G163            649 non-null   int64     
 187  G164            649 non-null   int64     
 188  G165            649 non-null   int64     
 189  G166            649 non-null   int64     
 190  G167            649 non-null   int64     
 191  G168            649 non-null   int64     
 192  G169            649 non-null   int64     
 193  G170            649 non-null   int64     
 194  G171            649 non-null   int64     
 195  G172            649 non-null   int64     
 196  G173            649 non-null   int64     
 197  G174            649 non-null   int64     
 198  G175            649 non-null   int64     
 199  G176            649 non-null   int64     
 200  G177            649 non-null   int64     
 201  G178            649 non-null   int64     
 202  G179            649 non-null   int64     
 203  G180            649 non-null   int64     
 204  G181            649 non-null   int64     
 205  G182            649 non-null   int64     
 206  G183            649 non-null   int64     
 207  G184            649 non-null   int64     
 208  G185            649 non-null   int64     
 209  G186            649 non-null   int64     
 210  G187            649 non-null   int64     
 211  G188            649 non-null   int64     
 212  G189            649 non-null   int64     
 213  G190            649 non-null   int64     
 214  G191            649 non-null   int64     
 215  G192            649 non-null   int64     
 216  G193            649 non-null   int64     
 217  G194            649 non-null   int64     
 218  G195            649 non-null   int64     
 219  G196            649 non-null   int64     
 220  G197            649 non-null   int64     
 221  G198            649 non-null   int64     
 222  G199            649 non-null   int64     
 223  G200            649 non-null   int64     
 224  G201            649 non-null   int64     
 225  G202            649 non-null   int64     
 226  G203            649 non-null   int64     
 227  G204            649 non-null   int64     
 228  G205            649 non-null   int64     
 229  G206            649 non-null   int64     
 230  G207            649 non-null   int64     
 231  G208            649 non-null   int64     
 232  G209            649 non-null   int64     
 233  G210            649 non-null   int64     
 234  G211            649 non-null   int64     
 235  G212            649 non-null   int64     
 236  G213            649 non-null   int64     
 237  G214            649 non-null   int64     
 238  G215            649 non-null   int64     
 239  G216            649 non-null   int64     
 240  G217            649 non-null   int64     
 241  G218            649 non-null   int64     
 242  G219            649 non-null   int64     
 243  G220            649 non-null   int64     
 244  G221            649 non-null   int64     
 245  G222            649 non-null   int64     
 246  G223            649 non-null   int64     
 247  G224            649 non-null   int64     
 248  G225            649 non-null   int64     
 249  G226            649 non-null   int64     
 250  G227            649 non-null   int64     
 251  G228            649 non-null   int64     
 252  G229            649 non-null   int64     
 253  G230            649 non-null   int64     
 254  G231            649 non-null   int64     
 255  G232            649 non-null   int64     
 256  G233            649 non-null   int64     
 257  G234            649 non-null   int64     
 258  G235            649 non-null   int64     
 259  G236            649 non-null   int64     
 260  G237            649 non-null   int64     
 261  G238            649 non-null   int64     
 262  G239            649 non-null   int64     
 263  G240            649 non-null   int64     
 264  G241            649 non-null   int64     
 265  G242            649 non-null   int64     
 266  G243            649 non-null   int64     
 267  G244            649 non-null   int64     
 268  G245            649 non-null   int64     
 269  G246            649 non-null   int64     
 270  G247            649 non-null   int64     
 271  G248            649 non-null   int64     
 272  G249            649 non-null   int64     
 273  G250            649 non-null   int64     
 274  G251            649 non-null   int64     
 275  G252            649 non-null   int64     
 276  G253            649 non-null   int64     
 277  G254            649 non-null   int64     
 278  G255            649 non-null   int64     
 279  G256            649 non-null   int64     
 280  G257            649 non-null   int64     
 281  G258            649 non-null   int64     
 282  G259            649 non-null   int64     
 283  G260            649 non-null   int64     
 284  G261            649 non-null   int64     
 285  G262            649 non-null   int64     
 286  G263            649 non-null   int64     
 287  G264            649 non-null   int64     
 288  G265            649 non-null   int64     
 289  G266            649 non-null   int64     
 290  G267            649 non-null   int64     
 291  G268            649 non-null   int64     
 292  G269            649 non-null   int64     
 293  G270            649 non-null   int64     
 294  G271            649 non-null   int64     
 295  G272            649 non-null   int64     
 296  G273            649 non-null   int64     
 297  G274            649 non-null   int64     
 298  G275            649 non-null   int64     
 299  G276            649 non-null   int64     
 300  G277            649 non-null   int64     
 301  G278            649 non-null   int64     
 302  G279            649 non-null   int64     
 303  G280            649 non-null   int64     
 304  G281            649 non-null   int64     
 305  G282            649 non-null   int64     
 306  G283            649 non-null   int64     
 307  G284            649 non-null   int64     
 308  G285            649 non-null   int64     
 309  G286            649 non-null   int64     
 310  G287            649 non-null   int64     
 311  G288            649 non-null   int64     
 312  G289            649 non-null   int64     
 313  G290            649 non-null   int64     
 314  G291            649 non-null   int64     
 315  G292            649 non-null   int64     
 316  G293            649 non-null   int64     
 317  G294            649 non-null   int64     
 318  G295            649 non-null   int64     
 319  G296            649 non-null   int64     
 320  G297            649 non-null   int64     
 321  G298            649 non-null   int64     
 322  G299            649 non-null   int64     
 323  G300            649 non-null   int64     
 324  G301            649 non-null   int64     
 325  G302            649 non-null   int64     
 326  G303            649 non-null   int64     
 327  G304            649 non-null   int64     
 328  G305            649 non-null   int64     
 329  G306            649 non-null   int64     
 330  G307            649 non-null   int64     
 331  G308            649 non-null   int64     
 332  G309            649 non-null   int64     
 333  G310            649 non-null   int64     
 334  G311            649 non-null   int64     
 335  G312            649 non-null   int64     
 336  G313            649 non-null   int64     
 337  G314            649 non-null   int64     
 338  G315            649 non-null   int64     
 339  G316            649 non-null   int64     
 340  G317            649 non-null   int64     
 341  G318            649 non-null   int64     
 342  G319            649 non-null   int64     
 343  G320            649 non-null   int64     
 344  G321            649 non-null   int64     
 345  G322            649 non-null   int64     
 346  G323            649 non-null   int64     
 347  G324            649 non-null   int64     
 348  G325            649 non-null   int64     
 349  G326            649 non-null   int64     
 350  G327            649 non-null   int64     
 351  G328            649 non-null   int64     
 352  G329            649 non-null   int64     
 353  G330            649 non-null   int64     
 354  G331            649 non-null   int64     
 355  G332            649 non-null   int64     
 356  G333            649 non-null   int64     
 357  G334            649 non-null   int64     
 358  G335            649 non-null   int64     
 359  G336            649 non-null   int64     
 360  G337            649 non-null   int64     
 361  G338            649 non-null   int64     
 362  G339            649 non-null   int64     
 363  G340            649 non-null   int64     
 364  G341            649 non-null   int64     
 365  G342            649 non-null   int64     
 366  G343            649 non-null   int64     
 367  G344            649 non-null   int64     
 368  G345            649 non-null   int64     
 369  G346            649 non-null   int64     
 370  G347            649 non-null   int64     
 371  G348            649 non-null   int64     
 372  G349            649 non-null   int64     
 373  G350            649 non-null   int64     
 374  G351            649 non-null   int64     
 375  G352            649 non-null   int64     
 376  G353            649 non-null   int64     
 377  G354            649 non-null   int64     
 378  G355            649 non-null   int64     
 379  G356            649 non-null   int64     
 380  G357            649 non-null   int64     
 381  G358            649 non-null   int64     
 382  G359            649 non-null   int64     
 383  G360            649 non-null   int64     
 384  G361            649 non-null   int64     
 385  G362            649 non-null   int64     
 386  G363            649 non-null   int64     
 387  G364            649 non-null   int64     
 388  G365            649 non-null   int64     
 389  G366            649 non-null   int64     
 390  G367            649 non-null   int64     
 391  G368            649 non-null   int64     
 392  G369            649 non-null   int64     
 393  G370            649 non-null   int64     
 394  G371            649 non-null   int64     
 395  G372            649 non-null   int64     
 396  G373            649 non-null   int64     
 397  G374            649 non-null   int64     
 398  G375            649 non-null   int64     
 399  G376            649 non-null   int64     
 400  G377            649 non-null   int64     
 401  G378            649 non-null   int64     
 402  G379            649 non-null   int64     
 403  G380            649 non-null   int64     
 404  G381            649 non-null   int64     
 405  G382            649 non-null   int64     
 406  G383            649 non-null   int64     
 407  G384            649 non-null   int64     
 408  G385            649 non-null   int64     
 409  G386            649 non-null   int64     
 410  G387            649 non-null   int64     
 411  G388            649 non-null   int64     
 412  G389            649 non-null   int64     
 413  G390            649 non-null   int64     
 414  G391            649 non-null   int64     
 415  G392            649 non-null   int64     
 416  G393            649 non-null   int64     
 417  G394            649 non-null   int64     
 418  G395            649 non-null   int64     
 419  G396            649 non-null   int64     
 420  G397            649 non-null   int64     
 421  G398            649 non-null   int64     
 422  G399            649 non-null   int64     
 423  G400            649 non-null
```

Logged out session ends in 0 seconds

To save and continue working [Sign In](#) or [Register](#)

notebook2da9e7a91a Draft saved

File Edit View Run Help

Code

```

    /  fedu      649 non-null int64
  8  Mjob      649 non-null object
  9  Fjob      649 non-null object
10  reason     649 non-null object
11  guardian   649 non-null object
12  traveltime 649 non-null int64
13  studytime   649 non-null int64
14  failures    649 non-null int64
15  schoolsup  649 non-null object
16  famsup    649 non-null object
17  paid       649 non-null object
18  activities 649 non-null object
19  nursery    649 non-null object
20  higher     649 non-null object
21  internet   649 non-null object
22  romantic   649 non-null object
23  famrel     649 non-null int64
24  freetime   649 non-null int64
25  goout     649 non-null int64
26  Dalc      649 non-null int64
27  Walc      649 non-null int64
28  health    649 non-null int64
29  absences   649 non-null int64
30  G1        649 non-null int64
31  G2        649 non-null int64
32  G3        649 non-null int64
dtypes: int64(16), object(17)
memory usage: 167.4+ KB

```

+ Code + Markdown

Console

Type here to search

Logged out session ends in 0 seconds

To save and continue working [Sign In](#) or [Register](#)

notebook2da9e7a91a

File Edit View Run Help

Code

```

    /  fftime    649 non-null int64
  24  goout    649 non-null int64
  25  Dalc     649 non-null int64
  26  Walc     649 non-null int64
  27  health   649 non-null int64
  28  absences 649 non-null int64
  30  G1       649 non-null int64
  31  G2       649 non-null int64
  32  G3       649 non-null int64
dtypes: int64(16), object(17)
memory usage: 167.4+ KB

```

+ Code + Markdown

df=df.replace("?",np.Nan)

df.head()

	college	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	—	4	3	4	1	1	3	4	0	11	11
1	GP	F	17	U	GT3	T	1	1	at_home	other	—	5	3	3	1	1	3	2	9	11	11
2	GP	F	15	U	LE3	T	1	1	at_home	other	—	4	3	2	2	3	3	6	12	13	12
3	GP	F	15	U	GT3	T	4	2	health	services	—	3	2	2	1	1	5	0	14	14	14
4	GP	F	16	U	GT3	T	3	3	other	other	—	4	3	2	1	2	5	0	11	13	13

5 rows × 33 columns

+ Code + Markdown

Console

Type here to search

Logged out session ends in 0 seconds

To save and continue working [Sign In](#) or [Register](#)

notebook2da9e7a91a

File Edit View Run Help

[12]: df.isnull().any().any()

[12]: False

[13]: df.isnull().sum()

```
college      0
sex          0
age          0
address      0
famsize      0
Pstatus      0
Medu         0
Fedu         0
Mjob         0
Fjob         0
reason        0
guardian     0
traveltime   0
studytime    0
failures     0
schoolsup    0
famsup       0
paid          0
activities   0
nursery      0
higher       0
internet     0
romantic     0
famrel       0
freetime     0
goout        0
Dalc          0
Walc          0
health        0
absences     0
G1            0
G2            0
G3            0
```

Console

Type here to search

My Quick Converter

Logged out session ends in 0 seconds

To save and continue working [Sign In](#) or [Register](#)

notebook2da9e7a91a Draft saved

File Edit View Run Help

[12]: df.isnull().any().any()

[12]: False

[13]: df.isnull().sum()

```
college      0
sex          0
age          0
address      0
famsize      0
Pstatus      0
Medu         0
Fedu         0
Mjob         0
Fjob         0
reason        0
guardian     0
traveltime   0
studytime    0
failures     0
schoolsup    0
famsup       0
paid          0
activities   0
nursery      0
higher       0
internet     0
romantic     0
famrel       0
freetime     0
goout        0
Dalc          0
Walc          0
health        0
absences     0
G1            0
G2            0
G3            0
```

Console

Type here to search

2:26 PM 14/03/2022

Logged out session ends in 0 seconds

To save and continue working [Sign In](#) or [Register](#)

notebook2da9e7a91a

File Edit View Run Help

[12]: df.isnull().any().any()

[12]: False

[13]: df.isnull().sum()

```
college      0
sex          0
age          0
address      0
famsize      0
Pstatus      0
Medu         0
Fedu         0
Mjob         0
Fjob         0
reason        0
guardian     0
traveltime   0
studytime    0
failures     0
schoolsup    0
famsup       0
paid          0
activities   0
nursery      0
higher       0
internet     0
romantic     0
famrel       0
freetime     0
goout        0
Dalc          0
Walc          0
health        0
absences     0
G1            0
G2            0
G3            0
```

Console

Type here to search

2:26 PM 14/03/2022

My Quick Converter | kaggle - Saferbrowser Yahoo Inc | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 2 seconds

To save and continue working Sign In or Register

notebook2da9e7a91a Draft saved

File Edit View Run Help

Code + Run All Draft Session (5m)

5 rows × 33 columns

[10]: df.mean()

```
/opt/conda/lib/python3.7/site-packages/ipykernel_launcher.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric_only=None') is deprecated; in a future version this will raise TypeError. Select only valid columns before calling the reduction.
    """Entry point for launching an IPython kernel.
```

```
[10]: age      16.744222
Medu      2.514638
Fedu      2.306626
traveltime 1.568567
studytime  1.930663
failures   0.221880
famrel     3.930663
freetime    3.180277
goout      3.184900
Dalc      1.502311
Walc      2.288431
health     3.536210
absences   3.659476
G1        11.399076
G2        11.578108
G3        11.906009
dtype: float64
```

Console

Type here to search

2:37 PM 14/03/2022

Logged out session ends in 2 seconds

To save and continue working Sign In or Register

notebook2da9e7a91a

File Edit View Run Help

Code + Run All Draft Session (6m)

```
/opt/conda/lib/python3.7/site-packages/ipykernel_launcher.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric_only=None') is deprecated; in a future version this will raise TypeError. Select only valid columns before calling the reduction.
    """Entry point for launching an IPython kernel.
```

```
[10]: age      16.744222
Medu      2.514638
Fedu      2.306626
traveltime 1.568567
studytime  1.930663
failures   0.221880
famrel     3.930663
freetime    3.180277
goout      3.184900
Dalc      1.502311
Walc      2.288431
health     3.536210
absences   3.659476
G1        11.399076
G2        11.578108
G3        11.906009
dtype: float64
```

[11]: avg\_age=df['age'].astype("float").mean()

```
avg_age
```

[11]: 16.7442218798151

+ Code + Markdown

Console

Type here to search

2:38 PM 14/03/2022

My Quick Converter | kaggle - Saferbrowser Yahoo Ind... | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 2 seconds To save and continue working Sign In or Register

notebook2da9e7a91a

File Edit View Run Help

Code

```
avg_age=df['age'].astype("float").mean()
avg_age
```

[11]: 16.7442218798151

df['age'].replace(np.NaN, avg\_age, inplace = True)
df['age']

[14]:

	age
0	18
1	17
2	15
3	15
4	16
..	
644	19
645	18
646	18
647	17
648	18

Name: age, Length: 649, dtype: int64

+ Code + Markdown

[ ]:

Console

Type here to search

Draft saved

Logged out session ends in 1 second To save and continue working Sign In or Register

notebook2da9e7a91a

File Edit View Run All Code

Code

```
1    17
2    15
3    15
4    16
..
644   19
645   18
646   18
647   17
648   18
Name: age, Length: 649, dtype: int64
```

[15]:

```
avg_G1=df['G1'].astype(float).mean()
df['G1'].replace(np.NaN, avg_G1, inplace=True)
df['G1']
```

[15]:

	G1
0	9
1	9
2	12
3	14
4	11
..	
644	10
645	15
646	11
647	10
648	10

Name: G1, Length: 649, dtype: int64

Console

Type here to search

Logged out session ends in 1 second

To save and continue working [Sign In](#) or [Register](#)

**notebook2da9e7a91a**

File Edit View Run Help

[5]: df.head()

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	...	4	3	4	1	1	3	4	0	11	11
1	GP	F	17	U	GT3	T	1	1	at_home	other	...	5	3	3	1	1	3	2	9	11	11
2	GP	F	15	U	LE3	T	1	1	at_home	other	...	4	3	2	2	3	3	6	12	13	12
3	GP	F	15	U	GT3	T	4	2	health	services	...	3	2	2	1	1	5	0	14	14	14
4	GP	F	16	U	GT3	T	3	3	other	other	...	4	3	2	1	2	5	0	11	13	13

5 rows × 33 columns

[6]: avg\_G2=df["G2"].astype(float).mean(axis= 0)  
print("Average of G2 : ",avg\_G2)  
df[ "G2" ].replace(np.nan, avg\_G2, inplace=True)

Average of G2 : 11.570107858243452

+ Code + Markdown

**Console**

Type here to search

My Quick Converter | kaggle - Saferbrowser Yahoo Inc | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 1 second

To save and continue working [Sign In](#) or [Register](#)

**notebook2da9e7a91a**

File Edit View Run Help

[15]: df[ "age" ].dtype

[15]: dtype('int64')

[16]: avg\_G3=df["G3"].astype(float).mean(axis= 0)  
print("Average of G3 : ",avg\_G3)  
df[ "G3" ].replace(np.nan, avg\_G3, inplace=True)

Average of G3 : 11.906009244992296

[17]: avg\_health=df[ "health" ].astype(float).mean(axis= 0)  
print("Average of health : ",avg\_health)  
df[ "health" ].replace(np.nan, avg\_health, inplace=True)

Average of health : 3.536209553158706

+ Code + Markdown

**Console**

Type here to search

Logged out session ends in 1 second

To save and continue working [Sign In](#) or [Register](#)

notebook2da9e7a91a Draft saved

File Edit View Run Help

Code

```
print("Average of health : ",avg_health)
df["health"].replace(np.nan, avg_health, inplace=True)
```

Average of health : 3.536209553158706

[18]: df['health'].value\_counts()

```
5 249
3 124
4 108
1 90
2 78
Name: health, dtype: int64
```

[19]: df['health'].value\_counts().idxmax()

[19]: 5

Console

Type here to search

Logged out session ends in 22 seconds

To save and continue working [Sign In](#) or [Register](#)

notebook2da9e7a91a Draft saved

File Edit View Run Help

Code

```
df['health'].value_counts().idxmax()
```

[19]: 5

[21]: df["health"].replace(np.nan, "5", inplace=True)
df.head()

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	traveltime	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3	
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	...	4	3	4	1	1	3	4	0	11	11
1	GP	F	17	U	GT3	T	1	1	at_home	other	...	5	3	3	1	1	3	2	9	11	11
2	GP	F	15	U	LE3	T	1	1	at_home	other	...	4	3	2	2	3	3	6	12	13	12
3	GP	F	15	U	GT3	T	4	2	health	services	...	3	2	2	1	1	5	0	14	14	14
4	GP	F	16	U	GT3	T	3	3	other	other	...	4	3	2	1	2	5	0	11	13	13

5 rows × 33 columns

Console

Type here to search

My Quick Converter | kaggle - Saferbrowser Yahoo! | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 2 seconds To save and continue working Sign In or Register

notebook2da9e7a91a

File Edit View Run Help Draft Session (2m)

[3]: df=pd.read\_csv("../input/student-performance-data-set/student-por.csv")  
print("read the csv file")

read the csv file

[6]: before\_rows=df.shape[0]  
df.dropna(subset=["Medu"], axis=0, inplace=True)  
after\_rows=df.shape[0]  
print("Number of dropped rows {}".format(before\_rows - after\_rows))  
df.reset\_index(drop=True, inplace=True)

Number of dropped rows 0

[8]: df.shape

[8]: (649, 33)

+ Code + Markdown

Console

Type here to search

[9]: df.dtypes

[9]: school object  
sex object  
age int64  
address object  
famsize object  
Pstatus object  
Medu int64  
Fedu int64  
Mjob object  
Fjob object  
reason object  
guardian object  
traveltime int64  
studytime int64  
failures int64  
schoolsup object  
famsup object  
paid object  
activities object  
nursery object  
higher object  
internet object  
romantic object  
famrel int64  
freetime int64  
goout int64  
Dalc int64  
Walc int64  
health int64  
absences int64  
G1 int64

Console

Type here to search

Kaggle Notebook Session

Logged out session ends in 2 seconds

notebook2da9e7a91a Draft saved

File Edit View Run Help

Code Draft Session (9m)

```
famrel int64
freetime int64
gout int64
Dalc int64
Walc int64
health int64
absences int64
G1 int64
G2 int64
G3 int64
dtype: object
```

[12]:

```
df[["age", "G3"]] = df[["age", "G3"]].astype("float")
df[["health"]] = df[["health"]].astype("int")
df[["G1"]] = df[["G1"]].astype("float")
df[["G3"]] = df[["G3"]].astype("float")
df.head()
```

[12]:

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel	freetime	gout	Dalc	Walc	health	absences	G1	G2	G3
0	GP	F	18.0	U	GT3	A	4	4	at_home	teacher	...	4	3	4	1	1	3	4	0.0	11	11.0
1	GP	F	17.0	U	GT3	T	1	1	at_home	other	...	5	3	3	1	1	3	2	9.0	11	11.0
2	GP	F	15.0	U	LE3	T	1	1	at_home	other	...	4	3	2	2	3	3	6	12.0	13	12.0
3	GP	F	15.0	U	GT3	T	4	2	health	services	...	3	2	2	1	1	5	0	14.0	14	14.0
4	GP	F	16.0	U	GT3	T	3	3	other	other	...	4	3	2	1	2	5	0	11.0	13	13.0

5 rows × 33 columns

Console

Type here to search

Logged out session ends in 1 second

notebook2da9e7a91a Draft

File Edit View Run Help

Code Draft Session (0m)

read the csv file

[4]:

```
df[["Medu"]] = 2 / df[["Medu"]]
df.rename(columns={"Medu": "MEDU"}, inplace=True)
df.head()
```

[4]:

	school	sex	age	address	famsize	Pstatus	MEDU	Fedu	Mjob	Fjob	...	famrel	freetime	gout	Dalc	Walc	health	absences	G1	G2	G3
0	GP	F	18	U	GT3	A	0.500000	4	at_home	teacher	...	4	3	4	1	1	3	4	0	11	11
1	GP	F	17	U	GT3	T	2.000000	1	at_home	other	...	5	3	3	1	1	3	2	9	11	11
2	GP	F	15	U	LE3	T	2.000000	1	at_home	other	...	4	3	2	2	3	3	6	12	13	12
3	GP	F	15	U	GT3	T	0.500000	2	health	services	...	3	2	2	1	1	5	0	14	14	14
4	GP	F	16	U	GT3	T	0.666667	3	other	other	...	4	3	2	1	2	5	0	11	13	13

5 rows × 33 columns

Console

Type here to search

Kaggle Notebook Interface (Screenshot 1)

Logged out session ends in 1 second

notebook2da9e7a91a

File Edit View Run Help

+ Run All Code Draft Session (3m)

[5]:

```
df['G1']=df['G1']+df['G1'].max()
df['G2']=df['G2']+df['G2'].max()
df.head()
```

5 rows × 33 columns

	school	sex	age	address	famsize	Pstatus	MEDU	Fedu	Mjob	Fjob	... famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3	
0	GP	F	18	U	GT3	A	0.500000	4	at_home	teacher	...	4	3	4	1	1	3	4	19	30	11
1	GP	F	17	U	GT3	T	2.000000	1	at_home	other	...	5	3	3	1	1	3	2	28	30	11
2	GP	F	15	U	LE3	T	2.000000	1	at_home	other	...	4	3	2	2	3	3	6	31	32	12
3	GP	F	15	U	GT3	T	0.500000	2	health	services	...	3	2	2	1	1	5	0	33	33	14
4	GP	F	16	U	GT3	T	0.666667	3	other	other	...	4	3	2	1	2	5	0	30	32	13

5 rows × 33 columns

+ Code + Markdown

Console

Type here to search

My Quick Converter

Logged out session ends in 1 second

notebook2da9e7a91a

File Edit View Run Help

+ Run All Code Draft Session (5m)

[7]:

```
df["G3"]=df["G3"].max()
df.head()
```

5 rows × 33 columns

	school	sex	age	address	famsize	Pstatus	MEDU	Fedu	Mjob	Fjob	... famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3	
0	GP	F	18	U	GT3	A	0.500000	4	at_home	teacher	...	4	3	4	1	1	3	4	19	30	0.578947
1	GP	F	17	U	GT3	T	2.000000	1	at_home	other	...	5	3	3	1	1	3	2	28	30	0.578947
2	GP	F	15	U	LE3	T	2.000000	1	at_home	other	...	4	3	2	2	3	3	6	31	32	0.631579
3	GP	F	15	U	GT3	T	0.500000	2	health	services	...	3	2	2	1	1	5	0	33	33	0.736842
4	GP	F	16	U	GT3	T	0.666667	3	other	other	...	4	3	2	1	2	5	0	30	32	0.684211

5 rows × 33 columns

+ Code + Markdown

Console

Type here to search

Logged out session ends in 1 second

To save and continue working [Sign In](#) or [Register](#)

notebook2da9e7a91a

File Edit View Run Help

+ Run All Code Draft Session (6m)

[8]: df.head()

	school	sex	age	address	famsize	Pstatus	MEDU	Fedu	Mjob	Fjob	... famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3	
0	GP	F	18	U	GT3	A	0.500000	4	at_home	teacher	...	4	3	4	1	1	3	4	19	30	0.578947
1	GP	F	17	U	GT3	T	2.000000	1	at_home	other	...	5	3	3	1	1	3	2	28	30	0.578947
2	GP	F	15	U	LE3	T	2.000000	1	at_home	other	...	4	3	2	2	3	3	6	31	32	0.631579
3	GP	F	15	U	GT3	T	0.500000	2	health	services	...	3	2	2	1	1	5	0	33	33	0.736842
4	GP	F	16	U	GT3	T	0.666667	3	other	other	...	4	3	2	1	2	5	0	30	32	0.684211

5 rows × 33 columns

[9]: df.to\_csv('Wrangled\_data.csv')

[+ Code] [+ Markdown]

Console

Type here to search

Logged out session ends in 1 second

To save and continue working [Sign In](#) or [Register](#)

notebook2da9e7a91a

File Edit View Run Help

+ Run All Code Draft Session (9m)

[9]: df.to\_csv('Wrangled\_data.csv')

[14]: df[ "Fedu" ]=df[ "Fedu" ].astype(int, copy=True)  
df.head()

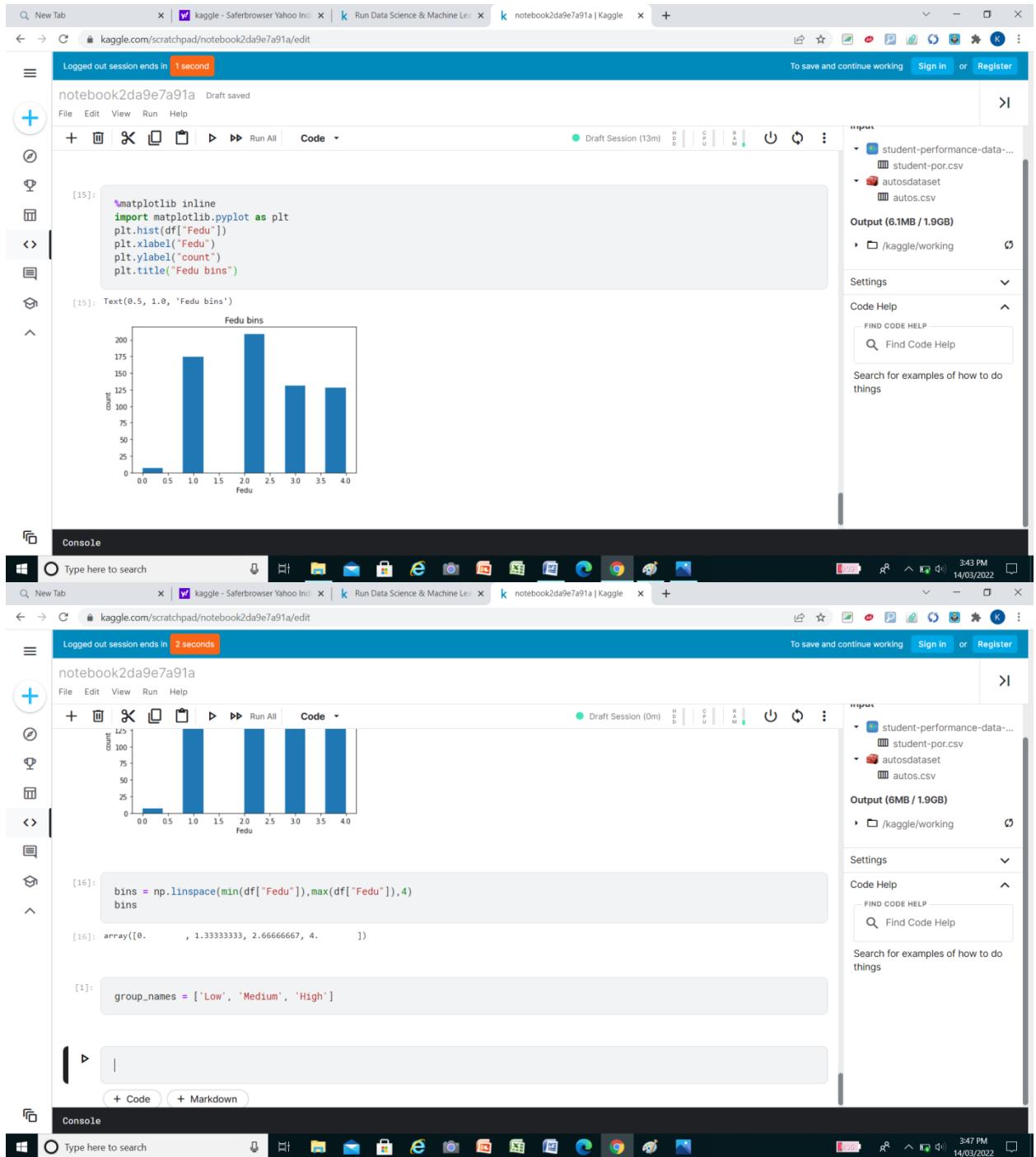
	school	sex	age	address	famsize	Pstatus	MEDU	Fedu	Mjob	Fjob	... famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3	
0	GP	F	18	U	GT3	A	0.500000	4	at_home	teacher	...	4	3	4	1	1	3	4	19	30	0.578947
1	GP	F	17	U	GT3	T	2.000000	1	at_home	other	...	5	3	3	1	1	3	2	28	30	0.578947
2	GP	F	15	U	LE3	T	2.000000	1	at_home	other	...	4	3	2	2	3	3	6	31	32	0.631579
3	GP	F	15	U	GT3	T	0.500000	2	health	services	...	3	2	2	1	1	5	0	33	33	0.736842
4	GP	F	16	U	GT3	T	0.666667	3	other	other	...	4	3	2	1	2	5	0	30	32	0.684211

5 rows × 33 columns

[+ Code] [+ Markdown]

Console

Type here to search



My Quick Converter | kaggle - Saferbrowser Yahoo! | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 2 seconds To save and continue working Sign In or Register

notebook2da9e7a91a

File Edit View Run Help

Code Draft Session (5m)

[6]:

```
bins = np.linspace(min(df["Fedu"]),max(df["Fedu"]),4)
bins
df[["Fedu-binned"]]=pd.cut(df[["Fedu"]], bins, labels=group_names, include_lowest=True )
df[["Fedu", "Fedu-binned"]].head(20)
```

[6]:

	Fedu	Fedu-binned
0	4	High
1	1	Low
2	1	Low
3	2	Medium
4	3	High
5	3	High
6	2	Medium
7	4	High
8	2	Medium
9	4	High
10	4	High
11	1	Low
12	4	High
13	3	High
14	2	Medium

Console

Type here to search

student-performance-data... student-por.csv autosdataset autos.csv

Output (6MB / 1.9GB) /kaggle/working

Settings

Code Help FIND CODE HELP Find Code Help

Search for examples of how to do things

3:52 PM 14/03/2022

Logged out session ends in 2 seconds To save and continue working Sign In or Register

notebook2da9e7a91a Draft saved

File Edit View Run Help

Code Draft Session (5m)

[6]:

	Fedu	Fedu-binned
0	4	High
1	1	Low
2	4	High
3	3	High
4	2	Medium
5	4	High
6	4	High
7	3	High
8	2	Medium
9	3	High

[7]:

```
df[["Fedu-binned"]].value_counts()
```

[7]:

```
High    259
Medium   209
Low     181
Name: Fedu-binned, dtype: int64
```

Console

Type here to search

student-performance-data... student-por.csv autosdataset autos.csv

Output (6MB / 1.9GB) /kaggle/working

Settings

Code Help FIND CODE HELP Find Code Help

Search for examples of how to do things

3:52 PM 14/03/2022

My Quick Converter | kaggle - Saferbrowser Yahoo! | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 2 seconds

To save and continue working [Sign In](#) or [Register](#)

notebook2da9e7a91a Draft saved

File Edit View Run Help

+ Run All Markdown

Name: Fedu-binned, dtype: int64

Draft Session (8m)

**input**

- student-performance-data-... student-por.csv
- autosdataset autos.csv

Output (6MB / 1.9GB)

/kaggle/working

Settings

Code Help

FIND CODE HELP

Find Code Help

Search for examples of how to do things

[8]:

```
%matplotlib inline
import matplotlib.pyplot as plt
plt.bar(group_names, df["Fedu-binned"].value_counts())
```

[8]: <BarContainer object of 3 artists>

%

+ Code + Markdown

Console

Type here to search

My Quick Converter | kaggle - Saferbrowser Yahoo! | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 2 seconds

To save and continue working [Sign In](#) or [Register](#)

notebook2da9e7a91a

File Edit View Run Help

+ Run All Markdown

Draft Session (11m)

**input**

- student-performance-data-... student-por.csv
- autosdataset autos.csv

Output (6MB / 1.9GB)

/kaggle/working

Settings

Code Help

FIND CODE HELP

Find Code Help

Search for examples of how to do things

[9]:

```
plt.hist(df[ "age" ], bins = 3)

plt.xlabel("age")
plt.ylabel("count")
plt.title("Age bins")
plt.show()
```

[9]: Age bins

%

Console

Type here to search

Logged out session ends in **2 seconds**

To save and continue working [Sign In](#) or [Register](#)

**notebook2da9e7a91a**

File Edit View Run Help

+ X D F Run All Markdown

Draft Session (14m)

input

- student-performance-data-... student-por.csv
- autosdataset autos.csv

Output (6MB / 1.9GB)

Settings

Code Help

FIND CODE HELP

Find Code Help

Search for examples of how to do things

[10]: df.columns

```
[10]: Index(['school', 'sex', 'age', 'address', 'famsize', 'Pstatus', 'Medu', 'Fedu', 'Mjob', 'Fjob', 'reason', 'guardian', 'traveltime', 'studystyle', 'failures', 'schoolsup', 'famsup', 'paid', 'activities', 'nursery', 'higher', 'internet', 'romantic', 'famrel', 'freetime', 'goout', 'Dalc', 'Walc', 'health', 'absences', 'G1', 'G2', 'G3', 'Fedu-binned'], dtype='object')
```

[11]: dummy\_variable\_1=pd.get\_dummies(df["school"])
dummy\_variable\_1.head()

	GP	MS
0	1	0
1	1	0
2	1	0
3	1	0
4	1	0

Console

Type here to search

408 PM 14/03/2022

Logged out session ends in **2 seconds**

To save and continue working [Sign In](#) or [Register](#)

**notebook2da9e7a91a** Draft saved

File Edit View Run Help

+ X D F Run All Code

Draft Session (3m)

input

- student-performance-data-... student-por.csv
- autosdataset autos.csv

Output (6MB / 1.9GB)

Settings

Code Help

FIND CODE HELP

Find Code Help

Search for examples of how to do things

[3]: GP MS

	GP	MS
0	1	0
1	1	0
2	1	0
3	1	0
4	1	0

[4]: dummy\_variable\_1.rename(columns={'school-primary':'secondary', 'school-primary':'primary'}, inplace=True)
dummy\_variable\_1.head()

	GP	MS
0	1	0
1	1	0
2	1	0
3	1	0
4	1	0

Console

Type here to search

408 PM 14/03/2022

My Quick Converter | kaggle - Saferbrowser Yahoo India | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 2 seconds

To save and continue working [Sign In](#) or [Register](#)

notebook2da9e7a91a

File Edit View Run Help

Code Draft Session (10m)

```
[5]: df=pd.concat([df, dummy_variable_1], axis=1)
df.drop('school', axis=1, inplace=True)
dummy_variable_2=pd.get_dummies(df['age'])
dummy_variable_2.rename(columns={'std':'age-std','number':'age-number'}, inplace=True)
dummy_variable_2.head()
```

	15	16	17	18	19	20	21	22
0	0	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0	0
2	1	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	0
4	0	1	0	0	0	0	0	0

+ Code + Markdown

Console

Type here to search

My Quick Converter | kaggle - Saferbrowser Yahoo India | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 2 seconds

To save and continue working [Sign In](#) or [Register](#)

notebook2da9e7a91a

File Edit View Run Help

Code Draft Session (12m)

```
[6]: df=pd.concat([df, dummy_variable_2], axis=1)
df.drop('age', axis=1, inplace=True)
df.head(10)
```

	sex	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	reason	guardian	...	GP	MS	15	16	17	18	19	20	21	22
0	F	U	GT3	A	4	4	at_home	teacher	course	mother	...	1	0	0	0	0	1	0	0	0	0
1	F	U	GT3	T	1	1	at_home	other	course	father	...	1	0	0	0	1	0	0	0	0	0
2	F	U	LE3	T	1	1	at_home	other	other	mother	...	1	0	1	0	0	0	0	0	0	0
3	F	U	GT3	T	4	2	health	services	home	mother	...	1	0	1	0	0	0	0	0	0	0
4	F	U	GT3	T	3	3	other	other	home	father	...	1	0	0	1	0	0	0	0	0	0
5	M	U	LE3	T	4	3	services	other	reputation	mother	...	1	0	0	1	0	0	0	0	0	0
6	M	U	LE3	T	2	2	other	other	home	mother	...	1	0	0	1	0	0	0	0	0	0
7	F	U	GT3	A	4	4	other	teacher	home	mother	...	1	0	0	0	1	0	0	0	0	0
8	M	U	LE3	A	3	2	services	other	home	mother	...	1	0	1	0	0	0	0	0	0	0
9	M	U	GT3	T	3	4	other	other	home	mother	...	1	0	1	0	0	0	0	0	0	0

10 rows × 41 columns

+ Code + Markdown

Console

Type here to search

Kaggle Notebook Session

notebook2da9e7a91a | Draft saved

File Edit View Run Help

`df.describe()`

	Medu	Fedu	traveltime	studytime	failures	famrel	freetime	goout	Dalc	Walc	...	GP	MS	15	16
count	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	...	649.000000	649.000000	649.000000	649.000000
mean	2.514638	2.306626	1.568567	1.930663	0.221880	3.930663	3.180277	3.184900	1.502311	2.280431	...	0.651772	0.348228	0.172573	0.272727
std	1.134552	1.099931	0.748660	0.829510	0.593235	0.955717	1.051093	1.175766	0.924834	1.284383	...	0.476776	0.476776	0.378169	0.445705
min	0.000000	0.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000	1.000000	1.000000	...	0.000000	0.000000	0.000000	0.000000
25%	2.000000	1.000000	1.000000	1.000000	0.000000	4.000000	3.000000	2.000000	1.000000	1.000000	...	0.000000	0.000000	0.000000	0.000000
50%	2.000000	2.000000	1.000000	2.000000	0.000000	4.000000	3.000000	3.000000	1.000000	2.000000	...	1.000000	0.000000	0.000000	0.000000
75%	4.000000	3.000000	2.000000	2.000000	0.000000	5.000000	4.000000	4.000000	2.000000	3.000000	...	1.000000	1.000000	0.000000	1.000000
max	4.000000	4.000000	4.000000	4.000000	3.000000	5.000000	5.000000	5.000000	5.000000	5.000000	...	1.000000	1.000000	1.000000	1.000000

8 rows × 25 columns

+ Code + Markdown

Console

Type here to search

My Quick Converter | kaggle - Saferbrowser Yahoo Inc. | Run Data Science & Machine Learning | notebook2da9e7a91a | Kaggle

Logged out session ends in 2 seconds

To save and continue working Sign In or Register

notebook2da9e7a91a | Draft saved

File Edit View Run Help

`df.describe()`

	famrel	freetime	goout	Dalc	Walc	...	GP	MS	15	16	17	18	19	20	21	22
0.000000	649.000000	649.000000	649.000000	649.000000	649.000000	...	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	649.000000	
93.0663	3.180277	3.184900	1.502311	2.280431	...	0.651772	0.348228	0.172573	0.272727	0.275809	0.215716	0.049307	0.009245	0.003082	0.001541	
95.5717	1.051093	1.175766	0.924834	1.284380	...	0.476776	0.476776	0.378169	0.445705	0.447266	0.411636	0.216674	0.095779	0.055470	0.039253	
0.000000	1.000000	1.000000	1.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	
0.000000	3.000000	2.000000	1.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	
0.000000	3.000000	3.000000	1.000000	2.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	
0.000000	4.000000	4.000000	2.000000	3.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	
0.000000	5.000000	5.000000	5.000000	5.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	

+ Code + Markdown

Console

Type here to search