abdulsaabir / **Pandas**

<> **Code** | ⊙ Issues | ⣿ Pull requests | ⊙ Actions | ⊞ Projects | ⚠ Security | ⬚ Insights | ⚙ Settings

abdulsaabir pandas lessons

5c688a4 · 1 hour ago | 🕘 History

Pandas / **lesson1.ipynb**   ↑ Top

Preview   Code   Blame   `5996 lines (5996 loc) · 283 KB`   Raw ⧉ ⭳ ✎ ▾

```
import pandas as pd
```

In [8]:
```python
# data frame
df = pd.read_csv('/home/saabir/Desktop/pandas/data/survey_results_public.csv')
```

In [10]:
```python
# when it reads this in it's going to read it in as a data frame so data frames
# frame is basically just rows and columns of data we can see what a data frame
# looks like but just by printing it out
```

In [23]:
```python
df
```

Out[23]:

	Respondent	MainBranch	Hobbyist	OpenSourcer	OpenSource	Employment	Country	Student	E(
0	1	I am a student who is learning to code	Yes	Never	The quality of OSS and closed source software ...	Not employed, and not looking for work	United Kingdom	No	Primary/elem
1	2	I am a student who is learning to code	No	Less than once per year	The quality of OSS and closed source software ...	Not employed, but looking for work	Bosnia and Herzegovina	Yes, full-time	Secondary (e.g. Am high scho
		I am not primarily a			The quality of OSS and				Bachelor's

2	3	primary a developer, but I write co...	Yes	Never	of OSS and closed source software ...	Employed full-time	Thailand	No	(BA, BS, E
3	4	I am a developer by profession	No	Never	The quality of OSS and closed source software ...	Employed full-time	United States	No	Bachelor's (BA, BS, E
4	5	I am a developer by profession	Yes	Once a month or more often	OSS is, on average, of HIGHER quality than pro...	Employed full-time	Ukraine	No	Bachelor's (BA, BS, E
...	...	...	...	...	...	...	...	...	
88878	88377	NaN	Yes	Less than once a month but more than once per ...	The quality of OSS and closed source software ...	Not employed, and not looking for work	Canada	No	Primary/elen
88879	88601	NaN	No	Never	The quality of OSS and closed source software ...	NaN	NaN	NaN	
88880	88802	NaN	No	Never	NaN	Employed full-time	NaN	NaN	
					OSS is, on	Independent			

88881	88816	NaN	No	Never	average, of HIGHER quality than pro...	independent contractor, freelancer, or self-em...	NaN	NaN	
88882	88863	NaN	Yes	Less than once per year	OSS is, on average, of HIGHER quality than pro...	Not employed, and not looking for work	Spain	Yes, full-time	Profe: degree (J

88883 rows × 85 columns

In [15]:
```python
# The df.shape attribute in Python typically refers to the shape or dimensions of a DataFrame
df.shape
```

Out[15]: (88883, 85)

In [19]:
```python
# The df.info() method provides information about the number of rows and columns in the DataFrame,
# as well as the data types of all the columns.
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 88883 entries, 0 to 88882
Data columns (total 85 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   Respondent      88883 non-null  int64
 1   MainBranch      88331 non-null  object
 2   Hobbyist        88883 non-null  object
 3   OpenSourcer     88883 non-null  object
 4   OpenSource      86842 non-null  object
 5   Employment      87181 non-null  object
 6   Country         88751 non-null  object
 7   Student         87014 non-null  object
 8   EdLevel         86390 non-null  object
 9   UndergradMajor  75614 non-null  object
 10  EduOther        84260 non-null  object
 11  OrgSize         71791 non-null  object
 12  DevType         81335 non-null  object
 13  YearsCode       87938 non-null  object
```

```
 13   YearsCode              87938 non-null   object
 14   Age1stCode             87634 non-null   object
 15   YearsCodePro           74331 non-null   object
 16   CareerSat              72847 non-null   object
 17   JobSat                 70988 non-null   object
 18   MgrIdiot               61159 non-null   object
 19   MgrMoney               61157 non-null   object
 20   MgrWant                61232 non-null   object
 21   JobSeek                80555 non-null   object
 22   LastHireDate           79854 non-null   object
 23   LastInt                67155 non-null   object
 24   FizzBuzz               71344 non-null   object
 25   JobFactors             79371 non-null   object
 26   ResumeUpdate           77877 non-null   object
 27   CurrencySymbol         71392 non-null   object
 28   CurrencyDesc           71392 non-null   object
 29   CompTotal              55945 non-null   float64
 30   CompFreq               63268 non-null   object
 31   ConvertedComp          55823 non-null   float64
 32   WorkWeekHrs            64503 non-null   float64
 33   WorkPlan               68914 non-null   object
 34   WorkChallenge          68141 non-null   object
 35   WorkRemote             70284 non-null   object
 36   WorkLoc                70055 non-null   object
 37   ImpSyn                 71779 non-null   object
 38   CodeRev                70390 non-null   object
 39   CodeRevHrs             49790 non-null   float64
 40   UnitTests              62668 non-null   object
 41   PurchaseHow            61108 non-null   object
 42   PurchaseWhat           62029 non-null   object
 43   LanguageWorkedWith     87569 non-null   object
 44   LanguageDesireNextYear 84088 non-null   object
 45   DatabaseWorkedWith     76026 non-null   object
 46   DatabaseDesireNextYear 69147 non-null   object
 47   PlatformWorkedWith     80714 non-null   object
 48   PlatformDesireNextYear 77443 non-null   object
 49   WebFrameWorkedWith     65022 non-null   object
 50   WebFrameDesireNextYear 62944 non-null   object
 51   MiscTechWorkedWith     59586 non-null   object
 52   MiscTechDesireNextYear 64511 non-null   object
 53   DevEnviron             87317 non-null   object
 54   OpSys                  87851 non-null   object
 55   Containers             85366 non-null   object
 56   BlockchainOrg          48175 non-null   object
 57   BlockchainIs           60165 non-null   object
```

```
 57   Stock+1n15              88185 non-null  object
 58   BetterLife              86269 non-null  object
 59   ITperson                87141 non-null  object
 60   OffOn                   86663 non-null  object
 61   SocialMedia             84437 non-null  object
 62   Extraversion            87305 non-null  object
 63   ScreenName              80486 non-null  object
 64   SOVisit1st              83877 non-null  object
 65   SOVisitFreq             88263 non-null  object
 66   SOVisitTo               88086 non-null  object
 67   SOFindAnswer            87816 non-null  object
 68   SOTimeSaved             86344 non-null  object
 69   SOHowMuchTime           68378 non-null  object
 70   SOAccount               87828 non-null  object
 71   SOPartFreq              74692 non-null  object
 72   SOJobs                  88066 non-null  object
 73   EntTeams                87841 non-null  object
 74   SOComm                  88131 non-null  object
 75   WelcomeChange           85855 non-null  object
 76   SONewContent            69560 non-null  object
 77   Age                     79210 non-null  float64
 78   Gender                  85406 non-null  object
 79   Trans                   83607 non-null  object
 80   Sexuality               76147 non-null  object
 81   Ethnicity               76668 non-null  object
 82   Dependents              83059 non-null  object
 83   SurveyLength            86984 non-null  object
 84   SurveyEase              87081 non-null  object
dtypes: float64(5), int64(1), object(79)
memory usage: 57.6+ MB
```

In [ ]:
```python
# we have 88883 enteries (rows) 85 columns , most parts Objects mean string
```

In [25]:
```python
pd.set_option('display.max_columns',85) # set the maximum columns to see (displayed ) one into 85
# remeber to go the first pd df variable you run case if you run under here you won't see any data in
```

In [ ]:
```python
#  the survey results schema CSV file that was included in our download gives the matching questions
# column names here so if we wanted to see what these column names here mean for this data then we ca
# (here) means the titles of the  data survey like => Respondent        MainBranch        Hobbyist
```

In [27]:
```python
df_schema = pd.read_csv('/home/saabir/Desktop/pandas/data/survey_results_schema.csv')
```

```
df_schema = pd.read_csv( /home/saabir/Desktop/pandas/data/survey_results_schema.csv )
```

In [28]:
```
df_schema
```

Out[28]:

	Column	QuestionText
0	Respondent	Randomized respondent ID number (not in order ...
1	MainBranch	Which of the following options best describes ...
2	Hobbyist	Do you code as a hobby?
3	OpenSourcer	How often do you contribute to open source?
4	OpenSource	How do you feel about the quality of open sour...
...	...	...
80	Sexuality	Which of the following do you currently identi...
81	Ethnicity	Which of the following do you identify as? Ple...
82	Dependents	Do you have any dependents (e.g., children, el...
83	SurveyLength	How do you feel about the length of the survey...
84	SurveyEase	How easy or difficult was this survey to compl...

85 rows × 2 columns

In [29]:
```
# let's set this up so that we can view 85 rows and then reprint
pd.set_option('display.max_row',85)
```

In [30]:
```
df_schema
```

Out[30]:

	Column	QuestionText
0	Respondent	Randomized respondent ID number (not in order ...
1	MainBranch	Which of the following options best describes ...

2	Hobbyist	Do you code as a hobby?
3	OpenSourcer	How often do you contribute to open source?
4	OpenSource	How do you feel about the quality of open sour...
5	Employment	Which of the following best describes your cur...
6	Country	In which country do you currently reside?
7	Student	Are you currently enrolled in a formal, degree...
8	EdLevel	Which of the following best describes the high...
9	UndergradMajor	What was your main or most important field of ...
10	EduOther	Which of the following types of non-degree edu...
11	OrgSize	Approximately how many people are employed by ...
12	DevType	Which of the following describe you? Please se...
13	YearsCode	Including any education, how many years have y...
14	Age1stCode	At what age did you write your first line of c...
15	YearsCodePro	How many years have you coded professionally (...
16	CareerSat	Overall, how satisfied are you with your caree...
17	JobSat	How satisfied are you with your current job? (...
18	MgrIdiot	How confident are you that your manager knows ...
19	MgrMoney	Do you believe that you need to be a manager t...
20	MgrWant	Do you want to become a manager yourself in th...
21	JobSeek	Which of the following best describes your cur...
22	LastHireDate	When was the last time that you took a job wit...
23	LastInt	In your most recent successful job interview (...
24	FizzBuzz	Have you ever been asked to solve FizzBuzz in ...

25	JobFactors	Imagine that you are deciding between two job ...
26	ResumeUpdate	Think back to the last time you updated your r...
27	CurrencySymbol	Which currency do you use day-to-day? If your ...
28	CurrencyDesc	Which currency do you use day-to-day? If your ...
29	CompTotal	What is your current total compensation (salar...
30	CompFreq	Is that compensation weekly, monthly, or yearly?
31	ConvertedComp	Salary converted to annual USD salaries using ...
32	WorkWeekHrs	On average, how many hours per week do you work?
33	WorkPlan	How structured or planned is your work?
34	WorkChallenge	Of these options, what are your greatest chall...
35	WorkRemote	How often do you work remotely?
36	WorkLoc	Where would you prefer to work?
37	ImpSyn	For the specific work you do, and the years of...
38	CodeRev	Do you review code as part of your work?
39	CodeRevHrs	On average, how many hours per week do you spe...
40	UnitTests	Does your company regularly employ unit tests ...
41	PurchaseHow	How does your company make decisions about pur...
42	PurchaseWhat	What level of influence do you, personally, ha...
43	LanguageWorkedWith	Which of the following programming, scripting,...
44	LanguageDesireNextYear	Which of the following programming, scripting,...
45	DatabaseWorkedWith	Which of the following database environments h...
46	DatabaseDesireNextYear	Which of the following database environments h...
47	PlatformWorkedWith	Which of the following platforms have you done...

48	PlatformDesireNextYear	Which of the following platforms have you done...
49	WebFrameWorkedWith	Which of the following web frameworks have you...
50	WebFrameDesireNextYear	Which of the following web frameworks have you...
51	MiscTechWorkedWith	Which of the following other frameworks, libra...
52	MiscTechDesireNextYear	Which of the following other frameworks, libra...
53	DevEnviron	Which development environment(s) do you use re...
54	OpSys	What is the primary operating system in which ...
55	Containers	How do you use containers (Docker, Open Contai...
56	BlockchainOrg	How is your organization thinking about or imp...
57	BlockchainIs	Blockchain / cryptocurrency technology is prim...
58	BetterLife	Do you think people born today will have a bet...
59	ITperson	Are you the "IT support person" for your family?
60	OffOn	Have you tried turning it off and on again?
61	SocialMedia	What social media site do you use the most?
62	Extraversion	Do you prefer online chat or IRL conversations?
63	ScreenName	What do you call it?
64	SOVisit1st	To the best of your memory, when did you first...
65	SOVisitFreq	How frequently would you say you visit Stack O...
66	SOVisitTo	I visit Stack Overflow to... (check all that a...
67	SOFindAnswer	On average, how many times a week do you find ...
68	SOTimeSaved	Think back to the last time you solved a codin...
69	SOHowMuchTime	About how much time did you save? If you're no...
70	SOAccount	Do you have a Stack Overflow account?
71	SOPartFreq	How frequently would you say you participate i

	SOPartFreq	How frequently would you say you participate ...
72	SOJobs	Have you ever used or visited Stack Overflow J...
73	EntTeams	Have you ever used Stack Overflow for Enterpri...
74	SOComm	Do you consider yourself a member of the Stack...
75	WelcomeChange	Compared to last year, how welcome do you feel...
76	SONewContent	Would you like to see any of the following on ...
77	Age	What is your age (in years)? If you prefer not...
78	Gender	Which of the following do you currently identi...
79	Trans	Do you identify as transgender?
80	Sexuality	Which of the following do you currently identi...
81	Ethnicity	Which of the following do you identify as? Ple...
82	Dependents	Do you have any dependents (e.g., children, el...
83	SurveyLength	How do you feel about the length of the survey...
84	SurveyEase	How easy or difficult was this survey to compl...

```
In [31]:   # to display the first few rows of a DataFrame use .head() method By default, it returns the first !
           # but If you want to display a different number of rows, you can pass the desired number as an argume
           df.head()
```

Out[31]:

	Respondent	MainBranch	Hobbyist	OpenSourcer	OpenSource	Employment	Country	Student	EdLev
0	1	I am a student who is learning to code	Yes	Never	The quality of OSS and closed source software ...	Not employed, and not looking for work	United Kingdom	No	Primary/elementa scho

		MainBranch	Hobbyist	OpenSourcer	OpenSource	Employment	Country	Student	EdLevel
1	2	I am a student who is learning to code	No	Less than once per year	The quality of OSS and closed source software ...	Not employed, but looking for work	Bosnia and Herzegovina	Yes, full-time	Secondary scho (e.g. America high school, G
2	3	I am not primarily a developer, but I write co...	Yes	Never	The quality of OSS and closed source software ...	Employed full-time	Thailand	No	Bachelor's degre (BA, BS, B.Eng etc
3	4	I am a developer by profession	No	Never	The quality of OSS and closed source software ...	Employed full-time	United States	No	Bachelor's degre (BA, BS, B.Eng etc
4	5	I am a developer by profession	Yes	Once a month or more often	OSS is, on average, of HIGHER quality than pro...	Employed full-time	Ukraine	No	Bachelor's degre (BA, BS, B.Eng etc

In [32]:
```python
df.head(10)
```

Out[32]:

	Respondent	**MainBranch**	**Hobbyist**	**OpenSourcer**	**OpenSource**	**Employment**	**Country**	**Student**	**EdLev**
		I am a student who			The quality of OSS and	Not employed,	United		Primary/elementa

0	1	student who is learning to code	Yes	Never	closed source software ...	and not looking for work	United Kingdom	No	Primary/elementa scho
1	2	I am a student who is learning to code	No	Less than once per year	The quality of OSS and closed source software ...	Not employed, but looking for work	Bosnia and Herzegovina	Yes, full-time	Secondary scho (e.g. America high school, G
2	3	I am not primarily a developer, but I write co...	Yes	Never	The quality of OSS and closed source software ...	Employed full-time	Thailand	No	Bachelor's degre (BA, BS, B.Eng etc
3	4	I am a developer by profession	No	Never	The quality of OSS and closed source software ...	Employed full-time	United States	No	Bachelor's degre (BA, BS, B.Eng etc
4	5	I am a developer by profession	Yes	Once a month or more often	OSS is, on average, of HIGHER quality than pro...	Employed full-time	Ukraine	No	Bachelor's degre (BA, BS, B.Eng etc
5	6	I am not primarily a developer	Yes	Never	The quality of OSS and closed	Employed	Canada	No	Bachelor's degre (BA, BS, B.Eng

5	6	developer, but I write co...	Yes	Never	closed source software ...	full-time	Canada	No	(BA, BS, B.Eng etc
6	7	I am a developer by profession	No	Never	The quality of OSS and closed source software ...	Independent contractor, freelancer, or self-em...	Ukraine	No	Bachelor's degre (BA, BS, B.Eng etc
7	8	I code primarily as a hobby	Yes	Less than once per year	OSS is, on average, of HIGHER quality than pro...	Not employed, but looking for work	India	NaN	Bachelor's degre (BA, BS, B.Eng etc
8	9	I am a developer by profession	Yes	Once a month or more often	The quality of OSS and closed source software ...	Employed full-time	New Zealand	No	Som college/universi study witho earning
9	10	I am a developer by profession	Yes	Once a month or more often	OSS is, on average, of HIGHER quality than pro...	Employed full-time	India	No	Master's degre (MA, MS, M.Eng MBA, etc

In [34]:
```
# to display the last few rows of a DataFrame use .tail() By default, it returns the last 5 rows of
# but you can specify a different number of rows to display .tail(number)
df.tail()
```

Out[34]:

Respondent	MainBranch	Hobbyist	OpenSourcer	OpenSource	Employment	Country	Student	EdLev

	Respondent	MainBranch	Hobbyist	OpenSourcer	OpenSource	Employment	Country	Student	Ed2...
88878	88377	NaN	Yes	Less than once a month but more than once per ...	The quality of OSS and closed source software ...	Not employed, and not looking for work	Canada	No	Primary/elementa school
88879	88601	NaN	No	Never	The quality of OSS and closed source software ...	NaN	NaN	NaN	N
88880	88802	NaN	No	Never	NaN	Employed full-time	NaN	NaN	N
88881	88816	NaN	No	Never	OSS is, on average, of HIGHER quality than pro...	Independent contractor, freelancer, or self-em...	NaN	NaN	N
88882	88863	NaN	Yes	Less than once per year	OSS is, on average, of HIGHER quality than pro...	Not employed, and not looking for work	Spain	Yes, full-time	Professio degree (JD, M et

In [ ]: