Adam_Syed_01_pandas_exercises

September 10, 2023

1 Pandas Exercises

- The goal of the exercises
  - Get comfortable with Pandas DataFrames
  - Establish workflows with Pandas
  - Deal with missing data
  - Make calculations on data
  - Export to Excel
- At each step, there are cells after the instructions where you should write your code. Just below that, there is a cell showing the output you should get.

1.0.1 Step 1

- import the Pandas library
- read the **gdp-00s.csv** file into a DataFrame **gdp00**
- ensure to set the **index_col** to 0
- Call **gdp00.head()**

[5]:
```
import pandas as pd
```

[6]:
```
gdp00 = pd.read_csv('gdp-00s.csv',index_col=0)
```

[ ]:
```

```

[7]:
```
gdp00.head()
```

[7]:

Country (or dependent territory)	2000	2001	2002	2003
Afghanistan	NaN	NaN	18768.0	20806.0
Albania	12857.0	14241.0	15116.0	16269.0
Algeria	251580.0	265033.0	284171.0	310706.0
Angola	37442.0	39906.0	46056.0	49439.0
Antigua and Barbuda	1305.0	1247.0	1271.0	1374.0

Country (or dependent territory)	2004	2005	2006	2007
Afghanistan	21522.0	24842.0	26978.0	31390.0
Albania	17639.0	19212.0	20971.0	22818.0

```
Algeria                          332978.0  363968.0  381471.0  404831.0
Angola                            56326.0   68755.0   85562.0  107684.0
Antigua and Barbuda                1463.0    1623.0    1875.0    2055.0

                                     2008      2009
Country (or dependent territory)
Afghanistan                       33242.0   40390.0
Albania                           25010.0   26045.0
Algeria                          422514.0  432672.0
Angola                           124967.0  128954.0
Antigua and Barbuda                2128.0    1914.0
```

1.0.2 Step 2

- read the **gdp-10s.csv** file into a DataFrame **gdp10**
- ensure to set the **index_col** to 0
- Call **gdp10.head()**

[8]:
```
gdp10 = pd.read_csv('gdp-10s.csv', index_col=0)
```

[ ]:
```
```

[9]:
```
gdp10.head()
```

[9]:
```
                                     2010      2011      2012      2013  \
Country (or dependent territory)
Afghanistan                       46304.0   50334.0   59945.0   63784.0
Albania                           28328.0   29655.0   30530.0   30604.0
Algeria                          471508.0  494947.0  497330.0  497988.0
Angola                           153871.0  162539.0  186124.0  199866.0
Antigua and Barbuda                1814.0    1816.0    1772.0    1720.0

                                     2014      2015      2016      2017  \
Country (or dependent territory)
Afghanistan                       69444.0   72056.0   70098.0   74712.0
Albania                           32529.0   33595.0   34736.0   37609.0
Algeria                          506135.0  477358.0  471382.0  478071.0
Angola                           220365.0  204604.0  204875.0  217987.0
Antigua and Barbuda                1762.0    1770.0    1883.0    1893.0

                                     2018      2019
Country (or dependent territory)
Afghanistan                       77416.0   81880.0
Albania                           40080.0   41709.0
Algeria                          496403.0  509307.0
Angola                           220542.0  222459.0
Antigua and Barbuda                2073.0    2181.0
```

1.0.3 Step 3

- concatenate the DataFrames **gdp00** and **gdp10** into **gdp**
- HINT: concatenate() exists in the pandas library.
- HINT: **axis=1** parameter
- Call **gdp.head()**

[12]:
```
gdp = pd.concat([gdp00, gdp10], axis=1)
```

[ ]:
```

```

[13]:
```
gdp.head()
```

[13]:

Country (or dependent territory)	2000	2001	2002	2003
Afghanistan	NaN	NaN	18768.0	20806.0
Albania	12857.0	14241.0	15116.0	16269.0
Algeria	251580.0	265033.0	284171.0	310706.0
Angola	37442.0	39906.0	46056.0	49439.0
Antigua and Barbuda	1305.0	1247.0	1271.0	1374.0

Country (or dependent territory)	2004	2005	2006	2007
Afghanistan	21522.0	24842.0	26978.0	31390.0
Albania	17639.0	19212.0	20971.0	22818.0
Algeria	332978.0	363968.0	381471.0	404831.0
Angola	56326.0	68755.0	85562.0	107684.0
Antigua and Barbuda	1463.0	1623.0	1875.0	2055.0

Country (or dependent territory)	2008	2009	2010	2011
Afghanistan	33242.0	40390.0	46304.0	50334.0
Albania	25010.0	26045.0	28328.0	29655.0
Algeria	422514.0	432672.0	471508.0	494947.0
Angola	124967.0	128954.0	153871.0	162539.0
Antigua and Barbuda	2128.0	1914.0	1814.0	1816.0

Country (or dependent territory)	2012	2013	2014	2015
Afghanistan	59945.0	63784.0	69444.0	72056.0
Albania	30530.0	30604.0	32529.0	33595.0
Algeria	497330.0	497988.0	506135.0	477358.0
Angola	186124.0	199866.0	220365.0	204604.0
Antigua and Barbuda	1772.0	1720.0	1762.0	1770.0

Country (or dependent territory)	2016	2017	2018	2019
Afghanistan	70098.0	74712.0	77416.0	81880.0

```
Albania                            34736.0   37609.0   40080.0   41709.0
Algeria                           471382.0  478071.0  496403.0  509307.0
Angola                            204875.0  217987.0  220542.0  222459.0
Antigua and Barbuda                 1883.0    1893.0    2073.0    2181.0
```

1.0.4 Step 4

- Drop all rows with **NaN**
- use argument **inplace=True**
- Call **gdp.head()**
- HINT: dropna() method on the DataFrame object

[14]:
```
gdp.dropna(inplace=True)
```

[ ]:
```

```

[15]:
```
gdp.head()
```

[15]:
```
                                    2000      2001      2002      2003  \
Country (or dependent territory)
Albania                          12857.0   14241.0   15116.0   16269.0
Algeria                         251580.0  265033.0  284171.0  310706.0
Angola                           37442.0   39906.0   46056.0   49439.0
Antigua and Barbuda               1305.0    1247.0    1271.0    1374.0
Argentina                       437650.0  427890.0  387126.0  429738.0

                                    2004      2005      2006      2007  \
Country (or dependent territory)
Albania                          17639.0   19212.0   20971.0   22818.0
Algeria                         332978.0  363968.0  381471.0  404831.0
Angola                           56326.0   68755.0   85562.0  107684.0
Antigua and Barbuda               1463.0    1623.0    1875.0    2055.0
Argentina                       481426.0  540900.0  602385.0  674119.0

                                    2008      2009      2010      2011  \
Country (or dependent territory)
Albania                          25010.0   26045.0   28328.0   29655.0
Algeria                         422514.0  432672.0  471508.0  494947.0
Angola                          124967.0  128954.0  153871.0  162539.0
Antigua and Barbuda               2128.0    1914.0    1814.0    1816.0
Argentina                       715230.0  678009.0  736718.0  797264.0

                                    2012      2013      2014      2015  \
Country (or dependent territory)
Albania                          30530.0   30604.0   32529.0   33595.0
Algeria                         497330.0  497988.0  506135.0  477358.0
Angola                          186124.0  199866.0  220365.0  204604.0
Antigua and Barbuda               1772.0    1720.0    1762.0    1770.0
```

```
Argentina                          819698.0  849616.0  839897.0  867177.0

                                       2016       2017       2018       2019
Country (or dependent territory)
Albania                             34736.0    37609.0    40080.0    41709.0
Algeria                            471382.0   478071.0   496403.0   509307.0
Angola                             204875.0   217987.0   220542.0   222459.0
Antigua and Barbuda                  1883.0     1893.0     2073.0     2181.0
Argentina                          885228.0  1039331.0  1036982.0  1033456.0
```

1.0.5 Step 5

- Calculate the GDP growth from 2000-2019 as a percentage in a column named **19y-pct**
- Call **gdp.head()**

[18]:
```
gdp['19y-pct'] = ((gdp['2019'] - gdp['2000']) / gdp['2000']) * 100
```

[ ]:
```
```

[19]:
```
gdp.head()
```

[19]:
```
                                    2000      2001      2002      2003  \
Country (or dependent territory)
Albania                          12857.0   14241.0   15116.0   16269.0
Algeria                         251580.0  265033.0  284171.0  310706.0
Angola                           37442.0   39906.0   46056.0   49439.0
Antigua and Barbuda               1305.0    1247.0    1271.0    1374.0
Argentina                       437650.0  427890.0  387126.0  429738.0

                                    2004      2005      2006      2007  \
Country (or dependent territory)
Albania                          17639.0   19212.0   20971.0   22818.0
Algeria                         332978.0  363968.0  381471.0  404831.0
Angola                           56326.0   68755.0   85562.0  107684.0
Antigua and Barbuda               1463.0    1623.0    1875.0    2055.0
Argentina                       481426.0  540900.0  602385.0  674119.0

                                    2008      2009  …       2011      2012  \
Country (or dependent territory)                    …
Albania                          25010.0   26045.0  …   29655.0   30530.0
Algeria                         422514.0  432672.0  …  494947.0  497330.0
Angola                          124967.0  128954.0  …  162539.0  186124.0
Antigua and Barbuda               2128.0    1914.0  …    1816.0    1772.0
Argentina                       715230.0  678009.0  …  797264.0  819698.0

                                    2013      2014      2015      2016  \
Country (or dependent territory)
Albania                          30604.0   32529.0   33595.0   34736.0
```

```
Algeria                           497988.0  506135.0  477358.0  471382.0
Angola                            199866.0  220365.0  204604.0  204875.0
Antigua and Barbuda                 1720.0    1762.0    1770.0    1883.0
Argentina                         849616.0  839897.0  867177.0  885228.0

                                       2017       2018       2019     19y-pct
Country (or dependent territory)
Albania                             37609.0    40080.0    41709.0  224.406938
Algeria                            478071.0   496403.0   509307.0  102.443358
Angola                             217987.0   220542.0   222459.0  494.142941
Antigua and Barbuda                  1893.0     2073.0     2181.0   67.126437
Argentina                         1039331.0  1036982.0  1033456.0  136.137553

[5 rows x 21 columns]
```

1.0.6 Step 6

- Calculate 10y-pct from 2009-2019
- Head

[20]:
```
gdp['10y-pct'] = ((gdp['2019'] - gdp['2009']) / gdp['2009']) * 100
```

[ ]:
```

```

[21]:
```
gdp.head()
```

[21]:
```
                                    2000      2001      2002      2003  \
Country (or dependent territory)
Albania                          12857.0   14241.0   15116.0   16269.0
Algeria                         251580.0  265033.0  284171.0  310706.0
Angola                           37442.0   39906.0   46056.0   49439.0
Antigua and Barbuda               1305.0    1247.0    1271.0    1374.0
Argentina                       437650.0  427890.0  387126.0  429738.0

                                    2004      2005      2006      2007  \
Country (or dependent territory)
Albania                          17639.0   19212.0   20971.0   22818.0
Algeria                         332978.0  363968.0  381471.0  404831.0
Angola                           56326.0   68755.0   85562.0  107684.0
Antigua and Barbuda               1463.0    1623.0    1875.0    2055.0
Argentina                       481426.0  540900.0  602385.0  674119.0

                                    2008      2009  …      2012      2013  \
Country (or dependent territory)                    …
Albania                          25010.0   26045.0  …   30530.0   30604.0
Algeria                         422514.0  432672.0  …  497330.0  497988.0
Angola                          124967.0  128954.0  …  186124.0  199866.0
Antigua and Barbuda               2128.0    1914.0  …    1772.0    1720.0
```

```
Argentina                          715230.0  678009.0  …  819698.0  849616.0

                                       2014      2015      2016       2017  \
Country (or dependent territory)
Albania                             32529.0   33595.0   34736.0    37609.0
Algeria                            506135.0  477358.0  471382.0   478071.0
Angola                             220365.0  204604.0  204875.0   217987.0
Antigua and Barbuda                  1762.0    1770.0    1883.0     1893.0
Argentina                          839897.0  867177.0  885228.0  1039331.0

                                       2018       2019     19y-pct    10y-pct
Country (or dependent territory)
Albania                             40080.0    41709.0  224.406938  60.142062
Algeria                            496403.0   509307.0  102.443358  17.712031
Angola                             220542.0   222459.0  494.142941  72.510353
Antigua and Barbuda                  2073.0     2181.0   67.126437  13.949843
Argentina                         1036982.0  1033456.0  136.137553  52.425115

[5 rows x 22 columns]
```

1.0.7 Step 7

- List the 10 countries with highest GDP growth rate in the 19 years 2000-2019.
- HINT:
  - sort_value
  - ascending=False
  - head(10)

[24]:
```
top_10_gdp_growth = gdp.sort_values(by='19y-pct', ascending=False).head(10)
```

[ ]:
```

```

[25]:
```
top_10_gdp_growth
```

[25]:
```
                                       2000       2001       2002       2003  \
Country (or dependent territory)
Turkmenistan                        11563.0    14238.0    16737.0    19989.0
Ethiopia                            32835.0    36075.0    37227.0    37172.0
Maldives                             1416.0     1498.0     1614.0     2064.0
Vietnam                            159786.0   174695.0   189936.0   207945.0
China                             3698622.0  4096897.0  4538333.0  5091708.0
Armenia                              7002.0     7840.0     9139.0    10630.0
Angola                              37442.0    39906.0    46056.0    49439.0
Bhutan                               1505.0     1647.0     1832.0     2039.0
Rwanda                               4997.0     5546.0     6373.0     6644.0
Cambodia                            13254.0    14660.0    15865.0    17557.0

                                       2004       2005       2006       2007  \
```

Country (or dependent territory)				
Turkmenistan	23556.0	27484.0	31435.0	35840.0
Ethiopia	42675.0	49617.0	57043.0	65468.0
Maldives	2399.0	2275.0	2812.0	3180.0
Vietnam	230306.0	255657.0	281900.0	310034.0
China	5760129.0	6617286.0	7686835.0	9011953.0
Armenia	12067.0	14213.0	16583.0	19365.0
Angola	56326.0	68755.0	85562.0	107684.0
Bhutan	2236.0	2459.0	2711.0	3133.0
Rwanda	7335.0	8282.0	9323.0	10301.0
Cambodia	19906.0	23268.0	26567.0	30059.0

	2008	2009	…	2012	\
Country (or dependent territory)			…		
Turkmenistan	41932.0	44842.0	…	63516.0	
Ethiopia	74220.0	82293.0	…	112533.0	
Maldives	3654.0	3485.0	…	5438.0	
Vietnam	334014.0	354717.0	…	568401.0	
China	10070854.0	11080887.0	…	15137455.0	
Armenia	21117.0	18266.0	…	27009.0	
Angola	124967.0	128954.0	…	186124.0	
Bhutan	3538.0	3769.0	…	5194.0	
Rwanda	11675.0	12500.0	…	15780.0	
Cambodia	32700.0	33644.0	…	42401.0	

	2013	2014	2015	\
Country (or dependent territory)				
Turkmenistan	68363.0	73746.0	76192.0	
Ethiopia	122434.0	148487.0	167119.0	
Maldives	6140.0	6957.0	7628.0	
Vietnam	607018.0	660612.0	700257.0	
China	16277355.0	17200686.0	17880337.0	
Armenia	28500.0	29231.0	29167.0	
Angola	199866.0	220365.0	204604.0	
Bhutan	5462.0	5871.0	6591.0	
Rwanda	16762.0	19002.0	20620.0	
Cambodia	45764.0	48654.0	52598.0	

	2016	2017	2018	\
Country (or dependent territory)				
Turkmenistan	78516.0	81788.0	88905.0	
Ethiopia	191322.0	215094.0	237224.0	
Maldives	8279.0	8930.0	9774.0	
Vietnam	770872.0	851064.0	933163.0	
China	18701703.0	19814058.0	21659302.0	
Armenia	31429.0	35677.0	38442.0	
Angola	204875.0	217987.0	220542.0	

```
Bhutan                                7270.0      7934.0      8434.0
Rwanda                               22076.0     23665.0     26312.0
Cambodia                             57942.0     62838.0     69193.0

                                        2019     19y-pct     10y-pct
Country (or dependent territory)
Turkmenistan                         96231.0  732.232120  114.600152
Ethiopia                            263111.0  701.312624  219.724642
Maldives                             10511.0  642.302260  201.606887
Vietnam                            1016475.0  536.147723  186.559426
China                             23393004.0  532.478907  111.111295
Armenia                              42095.0  501.185376  130.455491
Angola                              222459.0  494.142941   72.510353
Bhutan                                8912.0  492.159468  136.455293
Rwanda                               29298.0  486.311787  134.384000
Cambodia                             75390.0  468.809416  124.081560

[10 rows x 22 columns]
```

1.0.8 Step 8

- List the 10 countries with highest GDP growth rate between 2000-2019.
- HINT:
  - sort_value
  - ascending=False
  - head(10)

[26]:
```
top_10_gdp_growth = gdp.sort_values(by='19y-pct', ascending=False).head(10)
```

[ ]:
```
```

[27]:
```
top_10_gdp_growth
```

[27]:
```
                                      2000       2001       2002       2003  \
Country (or dependent territory)
Turkmenistan                       11563.0    14238.0    16737.0    19989.0
Ethiopia                           32835.0    36075.0    37227.0    37172.0
Maldives                            1416.0     1498.0     1614.0     2064.0
Vietnam                           159786.0   174695.0   189936.0   207945.0
China                            3698622.0  4096897.0  4538333.0  5091708.0
Armenia                             7002.0     7840.0     9139.0    10630.0
Angola                             37442.0    39906.0    46056.0    49439.0
Bhutan                              1505.0     1647.0     1832.0     2039.0
Rwanda                              4997.0     5546.0     6373.0     6644.0
Cambodia                           13254.0    14660.0    15865.0    17557.0

                                      2004       2005       2006       2007  \
Country (or dependent territory)
```

Turkmenistan	23556.0	27484.0	31435.0	35840.0
Ethiopia	42675.0	49617.0	57043.0	65468.0
Maldives	2399.0	2275.0	2812.0	3180.0
Vietnam	230306.0	255657.0	281900.0	310034.0
China	5760129.0	6617286.0	7686835.0	9011953.0
Armenia	12067.0	14213.0	16583.0	19365.0
Angola	56326.0	68755.0	85562.0	107684.0
Bhutan	2236.0	2459.0	2711.0	3133.0
Rwanda	7335.0	8282.0	9323.0	10301.0
Cambodia	19906.0	23268.0	26567.0	30059.0

	2008	2009	…	2012	\
Country (or dependent territory)			…		
Turkmenistan	41932.0	44842.0	…	63516.0	
Ethiopia	74220.0	82293.0	…	112533.0	
Maldives	3654.0	3485.0	…	5438.0	
Vietnam	334014.0	354717.0	…	568401.0	
China	10070854.0	11080887.0	…	15137455.0	
Armenia	21117.0	18266.0	…	27009.0	
Angola	124967.0	128954.0	…	186124.0	
Bhutan	3538.0	3769.0	…	5194.0	
Rwanda	11675.0	12500.0	…	15780.0	
Cambodia	32700.0	33644.0	…	42401.0	

	2013	2014	2015	\
Country (or dependent territory)				
Turkmenistan	68363.0	73746.0	76192.0	
Ethiopia	122434.0	148487.0	167119.0	
Maldives	6140.0	6957.0	7628.0	
Vietnam	607018.0	660612.0	700257.0	
China	16277355.0	17200686.0	17880337.0	
Armenia	28500.0	29231.0	29167.0	
Angola	199866.0	220365.0	204604.0	
Bhutan	5462.0	5871.0	6591.0	
Rwanda	16762.0	19002.0	20620.0	
Cambodia	45764.0	48654.0	52598.0	

	2016	2017	2018	\
Country (or dependent territory)				
Turkmenistan	78516.0	81788.0	88905.0	
Ethiopia	191322.0	215094.0	237224.0	
Maldives	8279.0	8930.0	9774.0	
Vietnam	770872.0	851064.0	933163.0	
China	18701703.0	19814058.0	21659302.0	
Armenia	31429.0	35677.0	38442.0	
Angola	204875.0	217987.0	220542.0	
Bhutan	7270.0	7934.0	8434.0	

```
Rwanda                                  22076.0     23665.0     26312.0
Cambodia                                57942.0     62838.0     69193.0

                                           2019     19y-pct     10y-pct
Country (or dependent territory)
Turkmenistan                            96231.0  732.232120  114.600152
Ethiopia                               263111.0  701.312624  219.724642
Maldives                                10511.0  642.302260  201.606887
Vietnam                               1016475.0  536.147723  186.559426
China                                23393004.0  532.478907  111.111295
Armenia                                 42095.0  501.185376  130.455491
Angola                                 222459.0  494.142941   72.510353
Bhutan                                   8912.0  492.159468  136.455293
Rwanda                                  29298.0  486.311787  134.384000
Cambodia                                75390.0  468.809416  124.081560

[10 rows x 22 columns]
```

1.0.9 Step 9

- Export data to Excel
- Name the file **gdp.xlsx**

[29]:
```
gdp.to_excel('gdp.xlsx')
```

[ ]:
```

```

2 Done