```
Python 3 (ipykernel) O
                                         Widgets
                                                                                                                  Trusted
                Insert
                                                       ✓ ☐ Autosave interval (min): off
          MPG Cars
          Introduction:
          The following exercise utilizes data from UC Irvine Machine Learning Repository
          Step 1. Import the necessary libraries
 In [1]: import pandas as pd
          import numpy as np
          Step 2. Import the first dataset <u>cars1</u> and <u>cars2</u>.
          Step 3. Assign each to a variable called cars1 and cars2
 In [3]: cars1 = pd.read_csv("https://raw.githubusercontent.com/guipsamora/pandas_exercises/master/05_Merge/Auto_MPG/cars1.csv")
 Out[3]:
                                                                                                    Unnamed:
                                                                                                              Unnamed:
                                                                                                                        Unnamed:
               mpg cylinders displacement horsepower weight acceleration model origin
                                                                                                                              12
                                                                                                                                        13
                                                                                   chevrolet
                                              130 3504
            0 18.0
                                    307
                                                                       70
                                                                12.0
                                                                                   chevelle
                                                                                                NaN
                                                                                                                   NaN
                                                                                                                             NaN
                                                                                                                                       NaN
                                                                                     malibu
                                                                                 buick
skylark 320
                                                                11.5
            1 15.0
                                    350
                                               165
                                                    3693
                                                                                                NaN
                                                                                                         NaN
                                                                                                                   NaN
                                                                                                                             NaN
                                                                                                                                       NaN
                                                                                   plymouth
                                              150 3436
                                    318
                                                                11.0
                                                                                                         NaN
            2 18.0
                          8
                                                                       70
                                                                                                NaN
                                                                                                                   NaN
                                                                                                                             NaN
                                                                                                                                       NaN
                                                                                  amc rebel
                                                                12.0
                                              150
                                                    3433
                                                                       70
                                                                                                NaN
                                                                                                         NaN
                                                                                                                             NaN
            3 16.0
                                    304
                                                                                                                   NaN
                                                                                                                                       NaN
                                                                                                                             NaN
                                                                                                                                       NaN
            4 17.0
                                    302
                                               140
                                                    3449
                                                                10.5
                                                                                                                   NaN
                                                                              1 ford torino
                                                                                                NaN
                                                                                                         NaN
                                               81 3012
                                                                17.6
           193 24.0
                                    200
                                                                                                                   NaN
                                                                                                                             NaN
                                                                                                                                       NaN
                                                                                   maverick
 In [4]: cars2 = pd.read_csv("https://raw.githubusercontent.com/guipsamora/pandas_exercises/master/05_Merge/Auto_MPG/cars2.csv")
          cars2
 Out[4]:
               mpg cylinders displacement horsepower weight acceleration model origin
                                                                                           car
                                    91
                                                53 1795
                                                                       76
            0 33.0
                                                                17.4
                                                                             3
                                                                                     honda civic
            1 20.0
                                    225
                                               100
                                                    3651
                                                                       76
                                                                17.7
                                                                                 dodge aspen se
            2 18.0
                                    250
                                               78 3574
                                                                21.0
                                                                       76
                                                                              1 ford granada ghia
            3 18.5
                          6
                                    250
                                               110 3645
                                                                16.2

    pontiac ventura sj

            4 17.5
                                    258
                                               95 3193
                                                                17.8
                                                                              1 amc pacer d/l
           195 27.0
                                    140
                                                86 2790
                                                                15.6
                                                                                  ford mustang gl
           196 44.0
                                               52 2130
                                     97
                                                                24.6
                                                                                      vw pickup
          197 32.0
                                    135
                                                    2295
                                                                11.6
                                                                                  dodge rampage
          198 28.0
                                               79 2625
                                    120
                                                                18.6
                                                                                      ford ranger
          199 31.0
                                               82 2720
                                                                19.4 82
                                    119
                                                                                      chevy s-10
          200 rows x 9 columns
          Step 4. Oops, it seems our first dataset has some unnamed blank columns, fix cars1
 In [5]: cars1 = cars1.dropna(axis=1)
          cars1
 Out[5]:
               mpg cylinders displacement horsepower weight acceleration model origin
                                               130 3504
            0 18.0
                                    307
                                                                12.0
                                                                       70

    chevrolet chevelle malibu

            1 15.0
                                    350
                                                    3693
                                                                       70
                                               165
                                                                11.5
                                                                                       buick skylark 320
            2 18.0
                                    318
                                               150
                                                    3436
                                                                11.0
                                                                       70
                                                                                       plymouth satellite
                                              150
                                                                12.0
            3 16.0
                                    304
                                                    3433
                                                                       70
                                                                                         amc rebel sst
                                                                       70
            4 17.0
                                    302
                                               140
                                                    3449
                                                                10.5
                                                                                            ford torino
                                    200
                                                81 3012
                                                                       76
           193 24.0
                                                                17.6
                                                                                         ford maverick
          194 22.5
                                    232
                                               90 3085
                                                                      76
                                                                17.6
                                                                                           amc hornet
          195 29.0
                                                                22.2
                                                                       76
                                     85
                                                52 2035
                                                                                      chevrolet chevette
          196 24.5
                                     98
                                                                22.1
                                                60
                                                    2164
                                                                       76
                                                                                       chevrolet woody
          197 29.0
                                                70 1937
                                                                14.2 76 2
                                                                                             vw rabbit
          198 rows × 9 columns
          Step 5. What is the number of observations in each dataset?
In [10]: cars1.shape[0],cars2.shape[0]
Out[10]: (198, 200)
          Step 6. Join cars1 and cars2 into a single DataFrame called cars
In [14]: car = pd.concat([cars1,cars2])
Out[14]:
               mpg cylinders displacement horsepower weight acceleration model origin
                                                    3504
            0 18.0
                                    307
                                               130
                                                                12.0
                                                                       70

    chevrolet chevelle malibu

            1 15.0
                          8
                                    350
                                                    3693
                                                                       70
                                               165
                                                                11.5
                                                                                       buick skylark 320
            2 18.0
                                    318
                                                    3436
                                                                11.0 70
                                                                                       plymouth satellite
                          8
                                                                       70
            3 16.0
                                    304
                                                    3433
                                               150
                                                                12.0
                                                                                         amc rebel sst
                                               140 3449
            4 17.0
                                    302
                                                                10.5 70 1
                                                                                            ford torino
           195 27.0
                                    140
                                                    2790
                                                                       82
                                                                                        ford mustang gl
                                                                15.6
                                    97
           196 44.0
                                                    2130
                                                                       82
                                                                              2
                          4
                                                                24.6
                                                                                            vw pickup
                                    135
           197 32.0
                                                    2295
                                                                11.6 82
                                                                                        dodge rampage
           198 28.0
                                    120
                                                79 2625
                                                                       82
                                                                18.6
                                                                                           ford ranger
                                    119
                                                                                           chevy s-10
           199 31.0
                                                82 2720
                                                                19.4 82
          398 rows x 9 columns
          Step 7. Oops, there is a column missing, called owners. Create a random number Series from 15,000 to
          73,000.
In [18]: owners = pd.Series(np.random.randint(15000, 73000, car.shape[0]))
          owners
Out[18]: 0
                 34813
                 19521
                 58311
                 32809
                 57738
                 17063
          394
                 22313
                 45045
          396
                 67461
                 18697
          397
          Length: 398, dtype: int32
          Step 8. Add the column owners to cars
In [19]: car["owners"] = owners
Out[19]:
               mpg cylinders displacement horsepower weight acceleration model origin
                                                                                                 car owners
                                                    3504
                                                                       70
            0 18.0
                                    307
                                               130
                                                                12.0
                                                                              1 chevrolet chevelle malibu 34813
            1 15.0
                                    350
                                                    3693
                                               165
                                                                11.5
                                                                       70
                                                                                       buick skylark 320
                                                                                                    19521
            2 18.0
                                    318
                                               150
                                                                                                    58311
                                                    3436
                                                                11.0
                                                                       70
                                                                                       plymouth satellite
            3 16.0
                                                    3433
                                                                       70
                                    304
                                                                                          amc rebel sst
                                    302
            4 17.0
                                               140 3449
                                                                10.5
                                                                       70
                                                                                            ford torino 57738
                                                                                        ford mustang gl 15479
           195 27.0
                                    140
                                                86 2790
           196 44.0
                                     97
                                                52 2130
                                                                       82
                                                                24.6
                                                                                            vw pickup
                                                                                                     19136
           197 32.0
                                    135
                                                    2295
                                                                                        dodge rampage 24516
                                                                11.6
           198 28.0
                                                    2625
                                                                                                     42315
                                                                       82
                                    120
                                                79
                                                                18.6
                                                                                           ford ranger
                                    119
                                                82 2720
                                                                       82
           199 31.0
                                                                19.4
                                                                                           chevy s-10 15689
          398 rows x 10 columns
In [20]: car.describe()
Out[20]:
                                                                          model
                            cylinders displacement
                                                     weight acceleration
                                                                                    origin
                                                                                               owners
                                                 398.000000
                                                            398.000000 398.000000 398.000000
                                                                                            398.000000
           count 398.000000 398.000000
                                       398.000000
                 23.514573
                            5.454774
                                       193.427136 2970.424623
                                                             15.568090
                                                                       76.010050
                                                                                  1.572864 44485.256281
                  7.815984
                            1.701004
                                       104.268683 846.841774
                                                              2.757689
                                                                        3.697627
                                                                                  0.802055 16712.430261
                            3.000000
                                                              8.000000
                                                                                  1.000000 15165.000000
                  9.000000
                                        68.000000 1613.000000
                                                                       70.000000
                 17.500000
                                                             13.825000
                                                                                   1.000000 30831.000000
                             4.000000
                                       104.250000 2223.750000
                                                                       73.000000
                             4.000000
                                       148.500000 2803.500000
                                                             15.500000
                                                                                   1.000000 44198.000000
                 23.000000
                                                                       76.000000
                             8.000000
                                       262.000000 3608.000000
                                                             17.175000
                                                                                  2.000000 59600.250000
                 29.000000
                                                                       79.000000
                                       455.000000 5140.000000
                                                                                  3.000000 72816.000000
                             8.000000
                                                             24.800000
                                                                      82.000000
In [23]: car.index
Out[23]: Int64Index([ 0, 1, 2, 3, 4, 5, 6, 7, 8, 9,
                      190, 191, 192, 193, 194, 195, 196, 197, 198, 199],
                     dtype='int64', length=398)
In [25]: car.info()
          <class 'pandas.core.frame.DataFrame'>
          Int64Index: 398 entries, 0 to 199
          Data columns (total 10 columns):
           # Column
                             Non-Null Count Dtype
                             398 non-null
                                             float64
             mpg
                             398 non-null
           1 cylinders
                                             int64
           2 displacement 398 non-null
                                             int64
             horsepower
                            398 non-null
                                             object
           4 weight
                             398 non-null
                                             int64
           5 acceleration 398 non-null
                                             float64
              model
                             398 non-null
                                             int64
           7 origin
                             398 non-null
                                             int64
                             398 non-null object
           8 car
                             398 non-null int32
           9 owners
          dtypes: float64(2), int32(1), int64(5), object(2)
          memory usage: 32.6+ KB
```

Jupyter Project 18 MPG Cars Last Checkpoint: a few seconds ago (autosaved)

Logout