

```

1  """ Udacity Programming for Data Scienarchive
2      Create: 30-May 2019 Last Mod: 16-June 2019
3      Python file name: bikeshare.py
4  Onport Library      """
5  import time
6  import pandas as pd
7  import numpy as np
8
9  """ Data definition """
10 CITY_DATA = { 'Chicago': 'chicago.csv',
11               'New York City': 'new_york_city.csv',
12               'Washington': 'washington.csv' }
13 CITIES     = ['Chicago', 'New York City', 'Washington']
14 MONTHS     = ['All', 'January', 'February', 'March', 'April', 'May',
15               • 'June']
16 DAYS       = ['All', 'Sunday', 'Monday', 'Tuesday', 'Wednesday',
17               • 'Thursday', 'Friday', 'Saturday' ]
18 sleeptime  = 2 #seconds
19 """ Progam Start """
20 def get_filters():
21     """ Input User city, month, and day to analyze.
22         Output (str) city, (str) month (str) day """
23     print('Explore some US bikeshare data')
24     # Input City
25     print('Which data of city like you: ', CITIES, '?')
26     city=input('City? ').title()
27     while city not in CITIES: city=input('Please again input City:
28     • ')
29     # Input Month
30     print('Which month? ', MONTHS, '?')
31     month=input('Month? ').title()
32     while month not in MONTHS: month=input('Please again input
33     • Month: ')
34     # Input Day
35     print('Which day ', DAYS, '?')
36     day=input('Day? ').title()
37     while day not in DAYS: day=input('Please again input Day:')
38
39     print('-'*48)
40     print('Now we will analyze US bikeshare data for')
41     print('City: ', city, 'in month: ', month, 'on days: ', day)
42     print('-'*48)
43     return city, month, day
44
45 def load_data(city, month, day):
46     """
47     Inpit data city, month, day from get_filters():
48     Output df - Pandas DataFrame """
49
50     df = pd.read_csv(CITY_DATA[city])

```

```

47 df['Start Time'] = pd.to_datetime(df['Start Time'])
48
49 df['month'] = df['Start Time'].dt.month
50 df['day_of_week'] = df['Start Time'].dt.weekday_name
51     # Filter by Month
52     if month != 'All':
53         months = ['January', 'February', 'March', 'April', 'May',
54     •         'June']
55         month = months.index(month) + 1
56         df = df[df['month'] == month]
57         # Filter by Day
58         if day != 'All':
59             df = df[df['day_of_week'] == day.title()]
60     return df
61
62 def time_stats(df,city,month,day):
63     """ Input data city, month, day from get_filters(); df from
64     • load_data()
65     • Output Displays statistics on the most frequent times of
66     • travel."""
67     print ( '\nCalculating The Most Frequent Times of Travel')
68     print("for City, Month, Day:  %s ,%s ,%s"%(city,month,day))
69     start_time = time.time()
70     # Display the most common month
71     common_month=df['month'].value_counts().head(1)
72     print("\ncommon month and count:  %s ."%(common_month))
73     # Display the most common day of week
74     common_day=df['day_of_week'].value_counts().head(1)
75     print("\ncommon day and count:  %s"%(common_day))
76     # Display the most common start hour
77     df['hour']=df['Start Time'].dt.hour
78     common_hour=df['hour'].value_counts().head(1)
79     print("\ncommon hour and cout:  %s"%(common_hour))
80
81     print("\nThis took %s seconds." % (time.time() - start_time))
82     print('-'*40)
83
84 def station_stats(df,city,month,day):
85     """
86     Input df from load_data() city,month,day from get_filters()
87     Output Displays statistics on the most popular stations and
88     • trip."""
89
90     print( '\nCalculating The Most Popular Stations and Trip')
91     print("for City, Month, Day:  %s ,%s ,%s"%(city,month,day))
92     start_time = time.time()
93     # Display most commonly used start station
94     common_start_station=df['Start Station'].value_counts().head(1)
95     print("\ncommon start station and count:
96     • %s"%(common_start_station))

```

```

92         # display most commonly used end station
93         common_end_station=df['End Station'].value_counts().head(1)
94         print("\ncommon start station and count:
    • %s"%(common_end_station))
95         # display most frequent combination of start station and
    • end station trip
96         df_group=df['Start Station']+df['End Station']
97         frequent_combination=df_group.value_counts().head(1)
98         print("\nthe most frequent
    • combination:%s"%(frequent_combination))
99
100         print("\nThis took %s seconds." % (time.time() - start_time))
101         print('-'*40)
102
103 def trip_duration_stats(df,city,month,day):
104     """
105     Input: df from load_data() city,month,day from
106     Output: Displays statistics on the total and average trip
    • duration."""
107     print('\nCalculating Trip Duration')
108     print("for City, Month, Day: %s ,%s ,%s"%(city,month,day))
109     start_time = time.time()
110     # Display total travel time
111     total_time=df['Trip Duration'].sum()
112     print("\ntotal travel time:%s's"%(total_time))
113     # Display mean travel time
114     mean_time=df['Trip Duration'].mean()
115     print("\nmean travel time:%s's"%(mean_time))
116
117     print("\nThis took %s seconds." % (time.time() - start_time))
118     label: end
119     print('-'*40)
120
121 def user_short (df,city,month,day):
122     """ Input df from load_data()
123     Output: Displays statistics on bikeshare users."""
124     # Data of with out gender burth information
125     print('\nCalculating User Stats')
126     print("for City, Month, Day: %s ,%s ,%s"%(city,month,day))
127     start_time = time.time()
128     # Display counts of user types
129     user_types=df['User Type'].value_counts()
130     print("\ncounts of user types:%s"%(user_types))
131     # Display earliest, most recent, and most common year of
    • birth
132     print ( '\nSory. Data about gender/birth are \nnot avabile for
    • '+ city)
133
134     print("\nThis took %s seconds." % (time.time() - start_time))
135     print('-'*40)

```

```

136
137 def user_stats(df,city,month,day):
138     """ Input df from load_data()
139         Output: Displays statistics on bikeshare users."""
140     print('\nCalculating User Stats')
141     print("for City, Month, Day:  %s ,%s ,%s"%(city,month,day))
142     start_time = time.time()
143     # Display counts of user types
144     user_types=df['User Type'].value_counts()
145     print("\ncounts of user types:%s"%(user_types))
146     # Display counts of gender
147     count_gender=df['Gender'].value_counts()
148     print("\ncounts of gender:%s"%(count_gender))
149     # Display earliest, most recent, and most common year of
    •     birth
150     earliest=df['Birth Year'].min()
151     most_recent=df['Birth Year'].max()
152     most_common=df['Birth Year'].mode()
153
154     print("\nthe earliest year:%s"%(earliest))
155     print("the most recent year:%s"%(most_recent))
156     print("the most common year:%s"%(most_common))
157
158     print("\nThis took %s seconds." % (time.time() - start_time))
159     print('-'*40)
160
161 def view_raw_data(df,city,month,day):
162     """
163     Input: df
164     Output: Displays the raw data
165     """
166     print('\nView Raw Data')
167
168     df = df.drop(['month', 'day_of_week'], axis = 1)
169     rowIndex = 0
170     seeData = input("\nWould you see the raw data of stats?\nPlease
    • write [y] Yes [n] No: ").lower()
171     while True:
172         if seeData == 'n':
173             return
174         if seeData == 'y':
175             print(df[rowIndex: rowIndex + 5])
176             rowIndex = rowIndex + 5
177             seeData = input("\nWould you see next five more
    • rows?\nPlease write [y] Yes [n] No: ").lower()
178
179 def main():
180     while True:
181         city, month, day = get_filters()
182         df = load_data(city, month, day)

```

```
183
184     time_stats(df,city,month,day)
185     time.sleep(sleeptime)
186     station_stats(df,city,month,day)
187     time.sleep(sleeptime)
188
189     trip_duration_stats(df,city,month,day)
190     time.sleep(sleeptime)
191     # sort solution city == 'Washington' have no gender data
192     if city == 'Washington':
193         user_short (df,city,month,day)
194     if city != 'Washington':
195         user_stats(df,city,month,day)
196
197     view_raw_data(df,city,month,day)
198     print('-'*40)
199
200     restart = input('\nWould you like to restart?\nEnter [y]
    • Yes [n] No: ')
201     if restart.lower() != 'y': break
202
203 if __name__ == "__main__":
204     main()
205
```