

```

1  import time
2  import pandas as pd
3  import numpy as np
4
5  CITY_DATA = { 'chicago': 'chicago.csv',
6                'new_york_city': 'new_york_city.csv',
7                'washington': 'washington.csv' }
8
9  def get_filters():
10     """
11     Asks user to specify a city, month, and day to analyze.
12
13     Returns:
14         (str) city - name of the city to analyze
15         (str) month - name of the month to filter by, or "all" to apply no month
16         filter
17         (str) day - name of the day of week to filter by, or "all" to apply no day
18         filter
19     """
20     print('Hello! Let\'s explore some US bikeshare data!')
21     # get user input for city (chicago, new york city, washington). HINT: Use a while
22     # loop to handle invalid inputs
23     city = input('ENTER THE CITY (chicago, new_york_city OR washington): ')
24     while city not in ['chicago', 'new_york_city', 'washington']:
25         city = input("CHOOSE BETWEEN chicago, new_york_city OR washington: ").lower()
26
27     # get user input for month (all, january, february, ... , june)
28     month = input('ENTER MONTH (all, january, february, ... , june): ').lower()
29     while month not in ['all', 'january', 'february', 'march', 'april', 'may', 'june']:
30         month = input('ENTER MONTH (all, january, february, ... , june) : ').lower()
31
32     # get user input for day of week (all, monday, tuesday, ... sunday)
33     day = input('ENTER DAY (all, monday, tuesday, ... sunday) : ').lower()
34     while day not in ['all', 'monday', 'tuesday', 'wednesday', 'thursday', 'friday',
35                       'saturday', 'sunday']:
36         day = input('ENTER DAY (all, monday, tuesday, ... sunday) : ').lower()
37
38     print('-'*40)
39     return city, month, day
40
41 def load_data(city, month, day):
42     """
43     Loads data for the specified city and filters by month and day if applicable.
44
45     Args:
46         (str) city - name of the city to analyze
47         (str) month - name of the month to filter by, or "all" to apply no month
48         filter
49         (str) day - name of the day of week to filter by, or "all" to apply no day
50         filter
51
52     Returns:
53         df - Pandas DataFrame containing city data filtered by month and day
54     """
55     df = pd.read_csv('{} .csv'.format(city))
56     df['Start Time'] = pd.to_datetime(df['Start Time'])
57     df['End Time'] = pd.to_datetime(df['End Time'])
58     df['month'] = df['Start Time'].dt.month
59     if month != 'all':
60         months = ['january', 'february', 'march', 'april', 'may', 'june']
61         month = months.index(month) + 1
62         df = df[df['month'] == month]
63     df['day_of_week'] = df['Start Time'].dt.day_name()
64     if day != 'all':
65         df = df[df['day_of_week'] == day.title()]
66     return df

```

```

64 def time_stats(df):
65     """Displays statistics on the most frequent times of travel."""
66
67     print('\nCalculating The Most Frequent Times of Travel...\n')
68     start_time = time.time()
69
70     # display the most common month
71     print("The most common month is: ", df['month'].value_counts().idxmax())
72
73     # display the most common day of week
74     print("The most common day is: ", df['day_of_week'].value_counts().idxmax())
75
76     # display the most common start hour
77     df['hour'] = df['Start Time'].dt.hour
78     print("The most common hour is: ", df['hour'].value_counts().idxmax())
79
80
81     print("\nThis took %s seconds." % (time.time() - start_time))
82     print('-'*40)
83
84
85 def station_stats(df):
86     """Displays statistics on the most popular stations and trip."""
87
88     print('\nCalculating The Most Popular Stations and Trip...\n')
89     start_time = time.time()
90
91     # display most commonly used start station
92     print("The most common start station is: ", df['Start Station'].value_counts().idxmax())
93
94     # display most commonly used end station
95     print("The most common end station is: ", df['End Station'].value_counts().idxmax())
96
97     # display most frequent combination of start station and end station trip
98     print("The most frequent combination of start station and end station trip")
99     most_common_start_and_end_stations = df.groupby(['Start Station', 'End Station']).size().nlargest(1)
100     print(most_common_start_and_end_stations)
101
102     print("\nThis took %s seconds." % (time.time() - start_time))
103     print('-'*40)
104
105
106 def trip_duration_stats(df):
107     """Displays statistics on the total and average trip duration."""
108
109     print('\nCalculating Trip Duration...\n')
110     start_time = time.time()
111
112     # display total travel time
113     total_duration = df['Trip Duration'].sum() / 3600.0
114     print("total travel time in hours is: ", total_duration)
115
116     # display mean travel time
117     mean_duration = df['Trip Duration'].mean() / 3600.0
118     print("mean travel time in hours is: ", mean_duration)
119
120
121     print("\nThis took %s seconds." % (time.time() - start_time))
122     print('-'*40)
123
124
125 def user_stats(df):
126     """Displays statistics on bikeshare users."""
127
128     print('\nCalculating User Stats...\n')
129     start_time = time.time()

```

```

130
131 # Display counts of user types
132 user_types = df['User Type'].value_counts()
133 print(user_types)
134
135 # Display counts of gender
136 try:
137     user_gender = df['Gender'].value_counts()
138     print("Gender is:", user_gender)
139
140
141     # Display earliest, most recent, and most common year of birth
142     earliest_year_of_birth = int(df['Birth Year'].min())
143     most_recent_year_of_birth = int(df['Birth Year'].max())
144     most_common_year_of_birth = int(df['Birth Year'].value_counts().idxmax())
145     print("The earliest year of birth is:", earliest_year_of_birth,
146           ", most recent one is:", most_recent_year_of_birth,
147           "and the most common one is: ", most_common_year_of_birth)
148 except:
149     print('No filter with gender allowed in Washington.')
150 print("\nThis took %s seconds." % (time.time() - start_time))
151 print('-'*40)
152
153 def display_data(df):
154     more_data = input("Would you like to view 5 rows of individual trip data? Enter\nyes or no? ").lower()
155     start_loc = 0
156     while more_data == 'yes':
157         print(df.iloc[start_loc:start_loc+5])
158         start_loc += 5
159         more_data = input("Do you wish to continue? Enter yes or no? ").lower()
160
161     return df
162
163 def main():
164     while True:
165         city, month, day = get_filters()
166         df = load_data(city, month, day)
167
168         time_stats(df)
169         station_stats(df)
170         trip_duration_stats(df)
171         user_stats(df)
172         display_data(df)
173
174         restart = input('\nWould you like to restart? Enter yes or no.\n')
175         if restart.lower() != 'yes':
176             break
177
178
179 if __name__ == "__main__":
180     main()
181

```