

```
1 # === DataFunctions.ipynb ===
2
3 # Dependencies
4 import pandas as pd
5
6
7 # Save path to data set in a variable
8 data_file = "Resources/dataSet.csv"
9
10
11 # Use Pandas to read data
12 data_file_pd = pd.read_csv(data_file)
13 data_file_pd.head()
14
15
16 # Display a statistical overview of the DataFrame
17 data_file_pd.describe()
18
19
20 # Reference a single column within a DataFrame
21 data_file_pd["Amount"].head()
22
23
24 # Reference multiple columns within a DataFrame
25 data_file_pd[["Amount", "Gender"]].head()
26
27
28 # The mean method averages the series
29 average = data_file_pd["Amount"].mean()
30 average
31
32
33 # The sum method adds every entry in the series
34 total = data_file_pd["Amount"].sum()
35 total
36
37
38 # The unique method shows every element of
39 # the series that appears only once
40 unique = data_file_pd["Last Name"].unique()
41 unique
42
43
44 # The value_counts method counts unique values in a column
45 count = data_file_pd["Gender"].value_counts()
```

```
46 count
47
48 # Calculations can also be performed on Series and
49 # added into DataFrames as new columns
50 thousands_of_dollars = data_file_pd["Amount"]/1000
51 data_file_pd["Thousands of Dollars"] = thousands_of_dollars
52
53 data_file_pd.head()
54
55
56
```