Assignment 6.1 Exploratory Data Analysis on Your Own Dataset

Name: Ballesteros, Angelo E.

Section: CPE22S2

Course Code and Description: CPE 311 - Computational Thinking with Python

Date Submitted and Performed: 6/24/24

Instructions:

- Choose your own dataset from public repositories.
- Link your chosen dataset in this spreadsheet: Dataset Spreadsheet
- Perform simple exploratory data analysis using descriptive statistics
- Employ the same steps as done in the hands-on activity and provide your own analysis of the dataset.
- Your submission must include your GitHub folder link that includes a customized readme file, Python Notebook Files, Dataset, and a simple presentation of your findings.

Note:

- Graphing is not required but may be done for additional points.

```
import pandas as pd


# upload csv file

filepath = '/content/valorant champions la.csv'
data = pd.read_csv(filepath, encoding='latin-1')

data
```

	Player	Team	Nationality	Kill	Death	K/D	KAST (%)	Prize ($)	Role	HS (%)	Rou Pla
0	Demon1	Evil Geniuses	United States	421	302	1.39	73	2,00,000	Flex	40	
1	Boostio	Evil Geniuses	United States	291	335	0.87	70	2,00,000	Flex	23	
2	jawgemo	Evil Geniuses	Cambodia	362	325	1.11	74	2,00,000	Flex	19	
3	Ethan	Evil Geniuses	United States	303	303	1.07	78	2,00,000	Flex	21	
4	C0M	Evil Geniuses	United States	330	294	1.12	74	2,00,000	Flex	21	
5	something	Paper Rex	Russia	269	245	1.10	72	80,000	Duelist	22	
6	Jinggg	Paper Rex	Singapore	285	279	1.02	72	80,000	Duelist	22	
7	f0rsakeN	Paper Rex	Indonesia	256	255	1.00	71	80,000	Flex	26	
8	d4v41	Paper Rex	Malaysia	243	227	1.07	77	80,000	Flex	30	
9	mindfreak	Paper Rex	Indonesia	214	220	0.97	72	80,000	Controller	30	
10	aspas	LOUD	Brazil	446	355	1.26	76	50,000	Duelist	25	
11	Less	LOUD	Brazil	416	344	1.21	73	50,000	Flex	24	
12	cauanzin	LOUD	Brazil	391	348	1.12	76	50,000	Initiator	25	
13	tuyz	LOUD	Brazil	304	317	0.96	75	50,000	Controller	29	
14	saadhak	LOUD	Argentina	310	355	0.87	73	50,000	Flex	19	
15	Boaster	FNATIC	United Kingdom	163	182	0.90	72	26,000	Controller	23	
16	Alfajer	FNATIC	Turkey	233	168	1.39	73	26,000	Sentinal	28	
17	Derke	FNATIC	Finland	206	180	1.14	66	26,000	Duelist	24	
18	Leo	FNATIC	Sweden	201	147	1.37	82	26,000	Initiator	27	
19	Chronical	FNATIC	Russia	156	176	0.89	72	26,000	Flex	23	
20	ZmjjKK	EDward Gaming	China	286	273	1.05	65	17,000	Duelist	18	
21	CHICHOO	EDward Gaming	China	245	236	1.04	72	17,000	Flex	16	
22	Smoggy	EDward Gaming	China	261	265	0.98	69	17,000	Flex	30	
23	nobody	EDward Gaming	China	248	266	0.93	72	17,000	Initiator	30	
24	Haodong	EDward Gaming	China	212	254	0.83	74	17,000	Controller	30	
25	MaKo	DRX	South Korea	183	189	0.97	73	17,000	Controller	31	
26	BuZz	DRX	South Korea	221	216	1.02	71	17,000	Duelist	22	
27	Zest	DRX	South Korea	197	201	0.98	67	17,000	Flex	35	
28	stax	DRX	South Korea	182	198	0.92	68	17,000	Initiator	33	
29	Rb	DRX	South Korea	163	191	0.85	65	17,000	Sentinal	23	
30	MrFaliN	FUT	Turkey	160	143	1.12	72	10,000	Flex	25	

Next steps: Generate code with data View recommended plots

```
data.columns # identify column names
```

```
Index(['Player', 'Team', 'Nationality', 'Kill', 'Death', 'K/D', 'KAST (%)',
       'Prize ($)', 'Role', 'HS (%)', 'Rounds Played', 'Rounds Win',
```

```
       'Rounds Lose', 'Rank'],
      dtype='object')
```

```
data.info() # identify data types
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 40 entries, 0 to 39
Data columns (total 14 columns):
 #   Column         Non-Null Count  Dtype  
---  ------         --------------  -----  
 0   Player         40 non-null     object 
 1   Team           40 non-null     object 
 2   Nationality    40 non-null     object 
 3   Kill           40 non-null     int64  
 4   Death          40 non-null     int64  
 5   K/D            40 non-null     float64
 6   KAST (%)       40 non-null     int64  
 7   Prize ($)      40 non-null     object 
 8   Role           40 non-null     object 
 9   HS (%)         40 non-null     int64  
 10  Rounds Played  40 non-null     int64  
 11  Rounds Win     40 non-null     int64  
 12  Rounds Lose    40 non-null     int64  
 13  Rank           40 non-null     object 
dtypes: float64(1), int64(7), object(6)
memory usage: 4.5+ KB
```

```
# display total number of records
```

```
data.shape[0]
```

```
40
```

```
# display first 20 records
```

```
data.head(20)
```

	Player	Team	Nationality	Kill	Death	K/D	KAST (%)	Prize ($)	Role	HS (%)	Rou Pla
0	Demon1	Evil Geniuses	United States	421	302	1.39	73	2,00,000	Flex	40	
1	Boostio	Evil Geniuses	United States	291	335	0.87	70	2,00,000	Flex	23	
2	jawgemo	Evil Geniuses	Cambodia	362	325	1.11	74	2,00,000	Flex	19	
3	Ethan	Evil Geniuses	United States	303	303	1.07	78	2,00,000	Flex	21	
4	C0M	Evil Geniuses	United States	330	294	1.12	74	2,00,000	Flex	21	
5	something	Paper Rex	Russia	269	245	1.10	72	80,000	Duelist	22	
6	Jinggg	Paper Rex	Singapore	285	279	1.02	72	80,000	Duelist	22	
7	f0rsakeN	Paper Rex	Indonesia	256	255	1.00	71	80,000	Flex	26	
8	d4v41	Paper Rex	Malaysia	243	227	1.07	77	80,000	Flex	30	
9	mindfreak	Paper Rex	Indonesia	214	220	0.97	72	80,000	Controller	30	
10	aspas	LOUD	Brazil	446	355	1.26	76	50,000	Duelist	25	
11	Less	LOUD	Brazil	416	344	1.21	73	50,000	Flex	24	
12	cauanzin	LOUD	Brazil	391	348	1.12	76	50,000	Initiator	25	
13	tuyz	LOUD	Brazil	304	317	0.96	75	50,000	Controller	29	

Next steps: Generate code with data View recommended plots

```
# display last 20 records

data.tail(20)
```

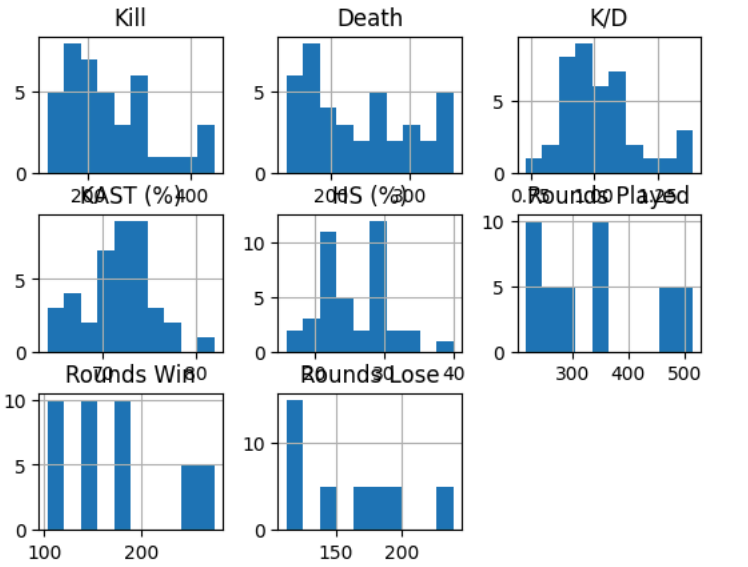
	Player	Team	Nationality	Kill	Death	K/D	KAST (%)	Prize ($)	Role	HS (%)	Rounds Played
20	ZmjjKK	EDward Gaming	China	286	273	1.05	65	17,000	Duelist	18	358
21	CHICHOO	EDward Gaming	China	245	236	1.04	72	17,000	Flex	16	358
22	Smoggy	EDward Gaming	China	261	265	0.98	69	17,000	Flex	30	358
23	nobody	EDward Gaming	China	248	266	0.93	72	17,000	Initiator	30	358
24	Haodong	EDward Gaming	China	212	254	0.83	74	17,000	Controller	30	358
25	MaKo	DRX	South Korea	183	189	0.97	73	17,000	Controller	31	283
26	BuZz	DRX	South Korea	221	216	1.02	71	17,000	Duelist	22	283
27	Zest	DRX	South Korea	197	201	0.98	67	17,000	Flex	35	283
28	stax	DRX	South Korea	182	198	0.92	68	17,000	Initiator	33	283
29	Rb	DRX	South Korea	163	191	0.85	65	17,000	Sentinal	23	283
30	MrFalin	FUT Esports	Turkey	160	143	1.12	72	10,000	Flex	25	216
31	AtaKaptan	FUT Esports	Turkey	136	154	0.88	70	10,000	Controller	30	216
32	qRaxs	FUT Esports	Turkey	145	155	0.94	72	10,000	Initiator	31	216
33	MOJJ	FUT Esports	Turkey	141	152	0.93	67	10,000	Flex	30	216
34	qw1	FUT	Turkey	151	163	0.93	64	10,000	Duelist	28	216

```
data.describe()
```

	Kill	Death	K/D	KAST (%)	HS (%)	Rounds Played	Rounds Win	
count	40.000000	40.000000	40.000000	40.000000	40.000000	40.000000	40.000000	40
mean	240.150000	233.825000	1.023500	71.575000	25.900000	335.750000	175.375000	160
std	84.153568	67.143126	0.152291	3.713368	5.062684	102.271321	60.259987	42
min	122.000000	143.000000	0.730000	64.000000	16.000000	216.000000	103.000000	113
25%	169.750000	175.000000	0.917500	70.000000	22.750000	259.250000	135.750000	122
50%	217.500000	223.500000	0.990000	72.000000	25.000000	320.500000	162.000000	155
75%	287.250000	282.750000	1.112500	73.000000	30.000000	383.250000	204.750000	185
max	446.000000	355.000000	1.390000	82.000000	40.000000	515.000000	276.000000	239

```
data.hist()
```

```
array([[<Axes: title={'center': 'Kill'}>,
        <Axes: title={'center': 'Death'}>,
        <Axes: title={'center': 'K/D'}>],
       [<Axes: title={'center': 'KAST (%)'}>,
        <Axes: title={'center': 'HS (%)'}>,
        <Axes: title={'center': 'Rounds Played'}>],
       [<Axes: title={'center': 'Rounds Win'}>,
        <Axes: title={'center': 'Rounds Lose'}>, <Axes: >]], dtype=object)
```




```
data.memory_usage()
```

```
Index            128
Player           320
Team             320
Nationality      320
Kill             320
Death            320
K/D              320
KAST (%)         320
Prize ($)        320
Role             320
HS (%)           320
Rounds Played    320
Rounds Win       320
Rounds Lose      320
Rank             320
dtype: int64
```

```
# print average of kills and deaths using numpy

import numpy as np

average_kill = np.mean(data['Kill'])
average_death = np.mean(data['Death'])

print(f"Average kill:", average_kill)
print(f"Average death:", average_death)
```

```
Average kill: 240.15
Average death: 233.825
```

```
# print median of rounds win and rounds lose using numpy

median_win = np.median(data['Rounds Win'])
median_lose = np.median(data['Rounds Lose'])

print(f"Median of rounds win:", median_win)
print(f"Median of rounds lose:", median_lose)
```

```
Median of rounds win: 162.0
Median of rounds lose: 155.0
```

```
# print the standard deviation of K/D using numpy

std_kd = np.std(data['K/D'])

print(f"Standard deviation of K/D:", std_kd)
```

```
Standard deviation of K/D: 0.1503753636737082
```

```
# print the average headshot using numpy

import numpy as np

average_headshot = np.mean(data['HS (%)'])

print(f"Average headshot percentage:", average_headshot)
```

```
Average headshot percentage: 25.9
```

```
# rename Nationality to Country

data.rename(columns = {'Nationality' : 'Country'})
```

	Player	Team	Country	Kill	Death	K/D	KAST (%)	Prize ($)	Role	HS (%)	Rounds Played
0	Demon1	Evil Geniuses	United States	421	302	1.39	73	2,00,000	Flex	40	459
1	Boostio	Evil Geniuses	United States	291	335	0.87	70	2,00,000	Flex	23	459
2	jawgemo	Evil Geniuses	Cambodia	362	325	1.11	74	2,00,000	Flex	19	459
3	Ethan	Evil Geniuses	United States	303	303	1.07	78	2,00,000	Flex	21	459
4	C0M	Evil Geniuses	United States	330	294	1.12	74	2,00,000	Flex	21	459
5	something	Paper Rex	Russia	269	245	1.10	72	80,000	Duelist	22	358
6	Jinggg	Paper Rex	Singapore	285	279	1.02	72	80,000	Duelist	22	358
7	f0rsakeN	Paper Rex	Indonesia	256	255	1.00	71	80,000	Flex	26	358
8	d4v41	Paper Rex	Malaysia	243	227	1.07	77	80,000	Flex	30	358
9	mindfreak	Paper Rex	Indonesia	214	220	0.97	72	80,000	Controller	30	358
10	aspas	LOUD	Brazil	446	355	1.26	76	50,000	Duelist	25	515
11	Less	LOUD	Brazil	416	344	1.21	73	50,000	Flex	24	515
12	cauanzin	LOUD	Brazil	391	348	1.12	76	50,000	Initiator	25	515
13	tuyz	LOUD	Brazil	304	317	0.96	75	50,000	Controller	29	515
14	saadhak	LOUD	Argentina	310	355	0.87	73	50,000	Flex	19	515
15	Boaster	FNATIC	United Kingdom	163	182	0.90	72	26,000	Controller	23	270
16	Alfajer	FNATIC	Turkey	233	168	1.39	73	26,000	Sentinal	28	270
17	Derke	FNATIC	Finland	206	180	1.14	66	26,000	Duelist	24	270
18	Leo	FNATIC	Sweden	201	147	1.37	82	26,000	Initiator	27	270
19	Chronical	FNATIC	Russia	156	176	0.89	72	26,000	Flex	23	270
20	ZmjjKK	EDward Gaming	China	286	273	1.05	65	17,000	Duelist	18	358
21	CHICHOO	EDward Gaming	China	245	236	1.04	72	17,000	Flex	16	358
22	Smoggy	EDward Gaming	China	261	265	0.98	69	17,000	Flex	30	358
23	nobody	EDward Gaming	China	248	266	0.93	72	17,000	Initiator	30	358
24	Haodong	EDward Gaming	China	212	254	0.83	74	17,000	Controller	30	358
25	MaKo	DRX	South Korea	183	189	0.97	73	17,000	Controller	31	283
26	BuZz	DRX	South Korea	221	216	1.02	71	17,000	Duelist	22	283
27	Zest	DRX	South Korea	197	201	0.98	67	17,000	Flex	35	283
28	stax	DRX	South Korea	182	198	0.92	68	17,000	Initiator	33	283

```
# delete Rank column

data.drop('Rank', axis = 1)
```

	Player	Team	Nationality	Kill	Death	K/D	KAST (%)	Prize ($)	Role	HS (%)	Rou Pla
0	Demon1	Evil Geniuses	United States	421	302	1.39	73	2,00,000	Flex	40	
1	Boostio	Evil Geniuses	United States	291	335	0.87	70	2,00,000	Flex	23	
2	jawgemo	Evil Geniuses	Cambodia	362	325	1.11	74	2,00,000	Flex	19	
3	Ethan	Evil Geniuses	United States	303	303	1.07	78	2,00,000	Flex	21	
4	C0M	Evil Geniuses	United States	330	294	1.12	74	2,00,000	Flex	21	
5	something	Paper Rex	Russia	269	245	1.10	72	80,000	Duelist	22	
6	Jinggg	Paper Rex	Singapore	285	279	1.02	72	80,000	Duelist	22	
7	f0rsakeN	Paper Rex	Indonesia	256	255	1.00	71	80,000	Flex	26	
8	d4v41	Paper Rex	Malaysia	243	227	1.07	77	80,000	Flex	30	
9	mindfreak	Paper Rex	Indonesia	214	220	0.97	72	80,000	Controller	30	
10	aspas	LOUD	Brazil	446	355	1.26	76	50,000	Duelist	25	
11	Less	LOUD	Brazil	416	344	1.21	73	50,000	Flex	24	
12	cauanzin	LOUD	Brazil	391	348	1.12	76	50,000	Initiator	25	
13	tuyz	LOUD	Brazil	304	317	0.96	75	50,000	Controller	29	
14	saadhak	LOUD	Argentina	310	355	0.87	73	50,000	Flex	19	
15	Boaster	FNATIC	United Kingdom	163	182	0.90	72	26,000	Controller	23	
16	Alfajer	FNATIC	Turkey	233	168	1.39	73	26,000	Sentinal	28	
17	Derke	FNATIC	Finland	206	180	1.14	66	26,000	Duelist	24	
18	Leo	FNATIC	Sweden	201	147	1.37	82	26,000	Initiator	27	
19	Chronical	FNATIC	Russia	156	176	0.89	72	26,000	Flex	23	
20	ZmjjKK	EDward Gaming	China	286	273	1.05	65	17,000	Duelist	18	
21	CHICHOO	EDward Gaming	China	245	236	1.04	72	17,000	Flex	16	
22	Smoggy	EDward Gaming	China	261	265	0.98	69	17,000	Flex	30	
23	nobody	EDward Gaming	China	248	266	0.93	72	17,000	Initiator	30	
24	Haodong	EDward Gaming	China	212	254	0.83	74	17,000	Controller	30	
25	MaKo	DRX	South Korea	183	189	0.97	73	17,000	Controller	31	
26	BuZz	DRX	South Korea	221	216	1.02	71	17,000	Duelist	22	
27	Zest	DRX	South Korea	197	201	0.98	67	17,000	Flex	35	
28	stax	DRX	South Korea	182	198	0.92	68	17,000	Initiator	33	
29	Rb	DRX	South Korea	163	191	0.85	65	17,000	Sentinal	23	
30	MrFalin	_ FUT	Turkey	160	143	1.12	72	10,000	Flex	25	

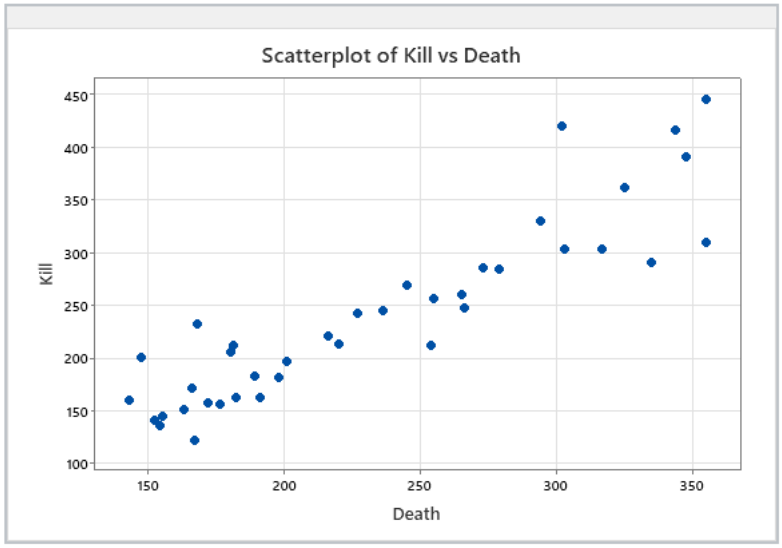
```
# return the mode of role in the tournament

data['Role'].mode()
```

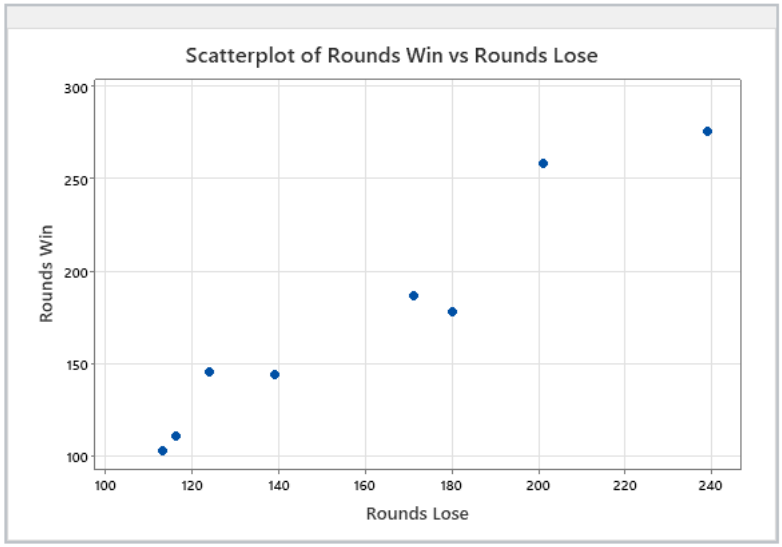
```
0    Flex
Name: Role, dtype: object
```

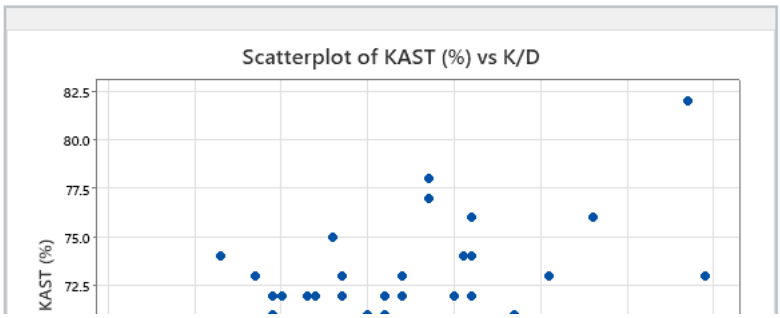
Descriptive Statistics: Kill, Death, K/D, KAST (%), HS (%), Rounds Played, Rounds Win, Rounds Lose, Rank

Statistics

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Kill	40	0	240.2	13.3	84.2	122.0	165.3	217.5	289.8	446.0
Death	40	0	233.8	10.6	67.1	143.0	173.0	223.5	290.3	355.0
K/D	40	0	1.0235	0.0241	0.1523	0.7300	0.9125	0.9900	1.1175	1.3900
KAST (%)	40	0	71.575	0.587	3.713	64.000	70.000	72.000	73.000	82.000
HS (%)	40	0	25.900	0.800	5.063	16.000	22.250	25.000	30.000	40.000
Rounds Played	40	0	335.8	16.2	102.3	216.0	237.8	320.5	433.8	515.0
Rounds Win	40	0	175.38	9.53	60.26	103.00	119.25	162.00	240.25	276.00
Rounds Lose	40	0	160.38	6.78	42.85	113.00	118.00	155.00	195.75	239.00
Rank	20	20	2.500	0.256	1.147	1.000	1.250	2.500	3.750	4.000







Analysis of the data

In the analysis, there were 40 player stats in the selected dataset and they are from different countries or regions. I have provided some findings of the data, including the mean, median, mode, and standard deviation of some selected variables such as Kill and Death. The majority of the role of this data is Flex role. Additionally, I have removed the Rank column because I think it is unnecessary to include it in the data since all of them are pro players. Lastly, I have used the provided statistics and graphs for the findings of the dataset.