

# UNIFIED MENTOR INTERN PROJECT 1

Data Visualization of Bird Strikes between 2000 – 2011

BY MADHUMITHA M

```
import pandas as pd
import matplotlib.pyplot as plt
```

```
# Load the dataset
```

```
file_path = r"C:\Users\VIS\Downloads\Bird Strikes data unified.xlsx"
```

```
data = pd.read_excel(file_path)
```

data

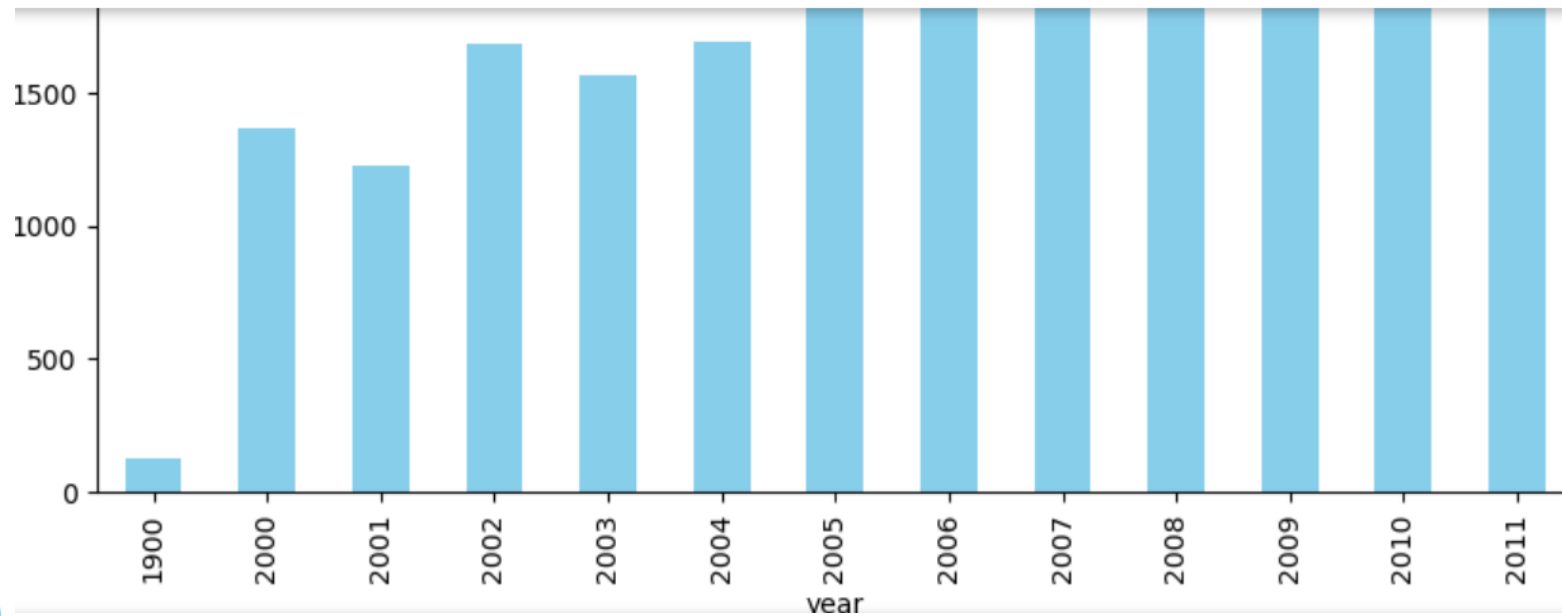
	Record ID	Aircraft: Type	Airport: Name	Altitude bin	Aircraft: Make/Model	Wildlife: Number struck	Wildlife: Number Struck Actual	Effect: Impact to flight	FlightDate	year	...	Remains of wildlife collected?	Remains of wildlife sent to Smithsonian	Wildlife: Size	Conditions: Sky	V
0	202152	Airplane	LAGUARDIA NY	> 1000 ft	B-737-400	Over 100	859	Engine Shut Down	2000-11-23	2000	...	False	False	Medium	No Cloud	U
1	208159	Airplane	DALLAS/FORT WORTH INTL ARPT	< 1000 ft	MD-80	Over 100	424	NaN	2001-07-25	2001	...	False	False	Small	Some Cloud	
2	207601	Airplane	LAKEFRONT AIRPORT	< 1000 ft	C-500	Over 100	261	NaN	2001-09-14	2001	...	False	False	Small	No Cloud	Ei
3	215953	Airplane	SEATTLE-TACOMA	< 1000 ft	B-737-400	Over 100	806	Precautionary Landing	2002-09-05	2002	...	True	False	Small	Some Cloud	Ei

## Data Visualization of Bird Strikes between 2000 - 2011

It looks like I uploaded an image named as data import visualization

However, to perform data visualization for bird strikes, we need the actual data file, preferably in CSV format.

```
ly_data = data.groupby('year').size()  
.figure(figsize=(10, 6))  
ly_data.plot(kind='bar', color='skyblue')  
.title('Yearly Bird Strikes in the US')  
.xlabel('year')  
.ylabel('Number of Bird Strikes')  
.show()
```



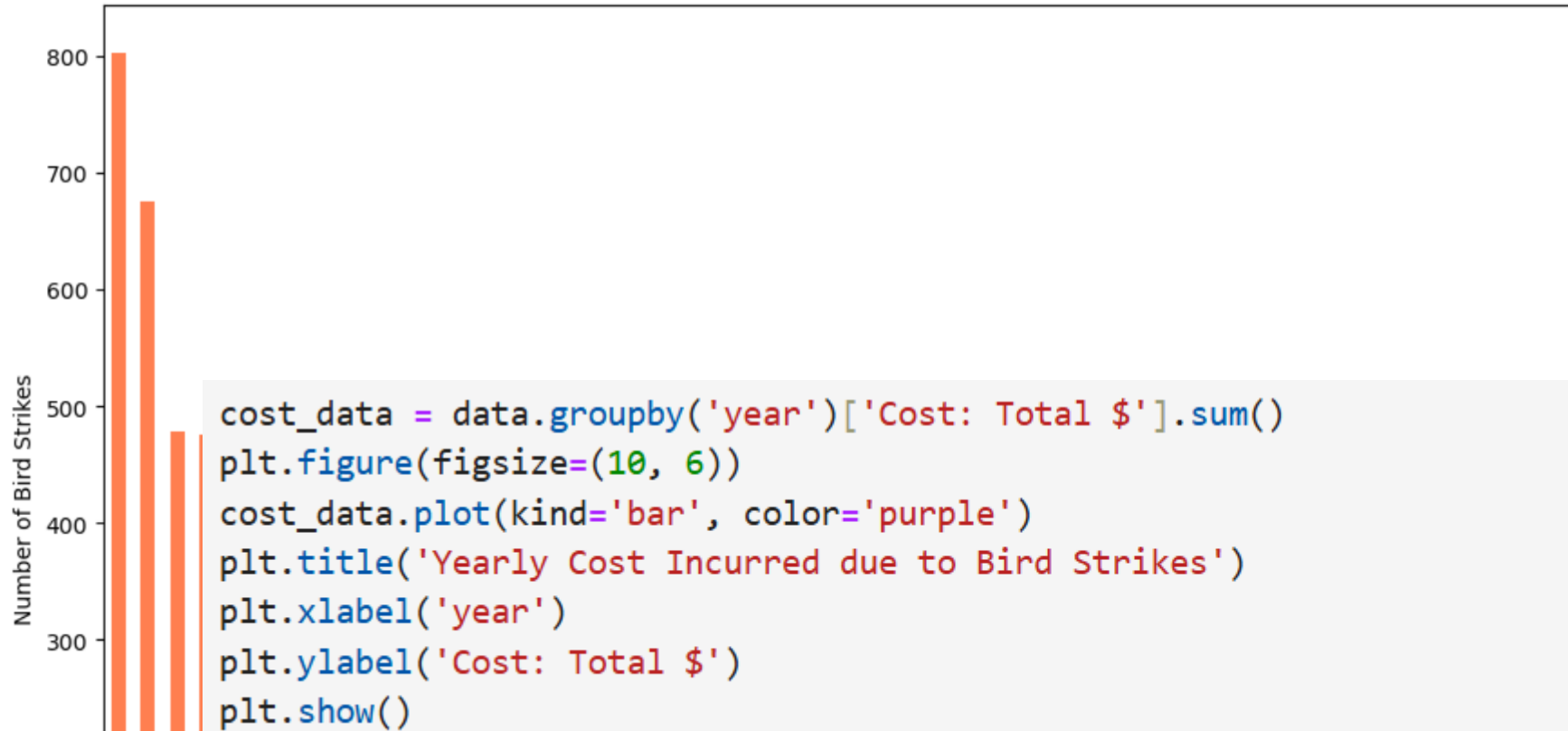
## YEARLY BRID SRIKES IN THE US

Above the visualization of the data yearly brid strikes in US



## Top US airline by bird strikes

A bar chart showing the top 10 US airlines by the number of bird strikes during the same period.



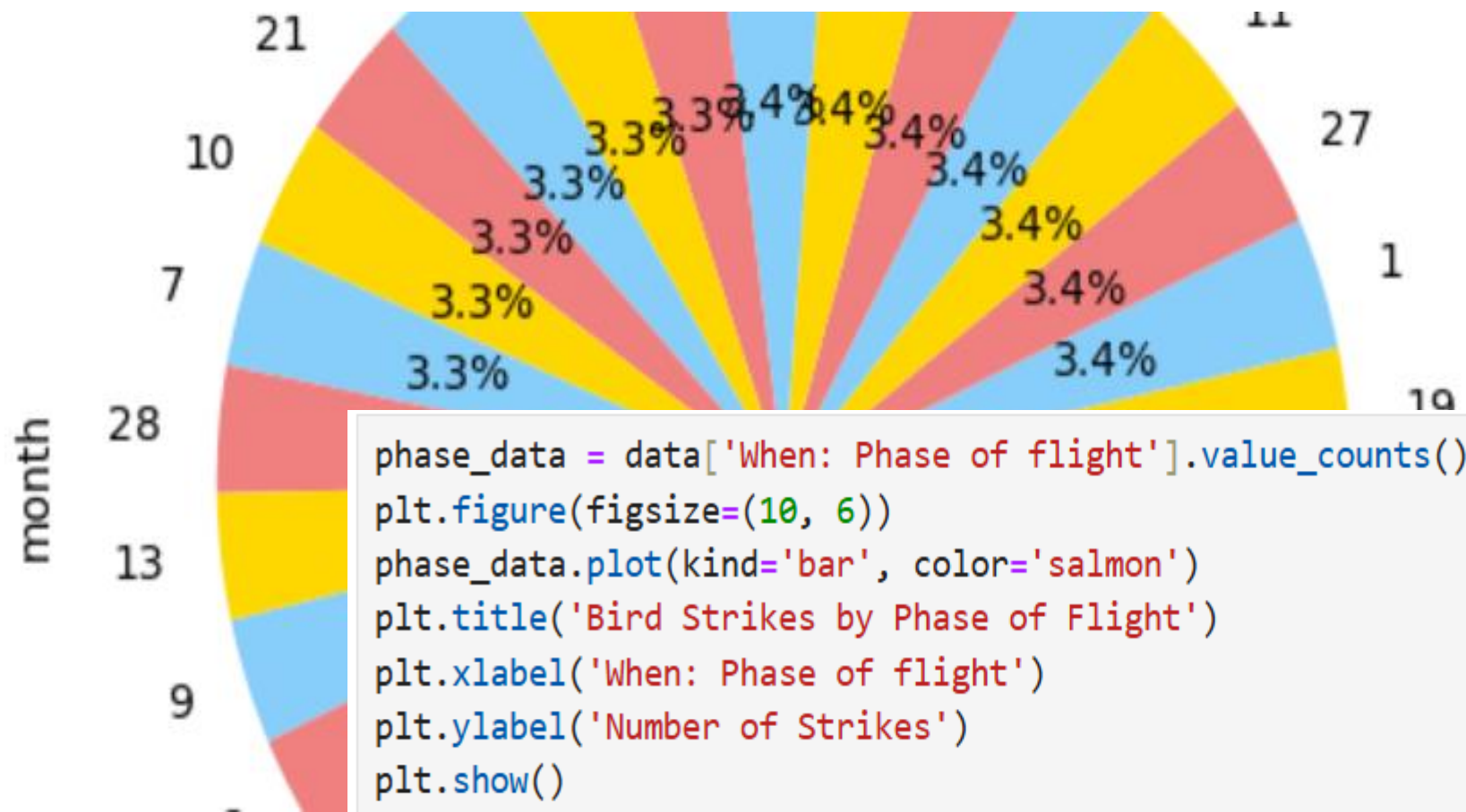
## Top 50 Airports by bird strikes

A bar chart showing the top 50 US airports by the number of bird strikes during the same period.



## Yearly cost incurred due to bird strikes

A bar chart showing the yearly cost incurred due to bird strikes top



## Bird strikes by time of day

A plot chart showing the month of bird strikes by time of day

Number of Strikes



```
effect_data = data['Effect: Impact to flight'].value_counts()
plt.figure(figsize=(10, 6))
effect_data.plot(kind='bar', color='plum')
plt.title('Effect of Bird Strikes on Flights')
plt.xlabel('Effect: Impact to flight')
plt.ylabel('Number of Strikes')
plt.show()
```

## Bird strikes by phase of light

A plot chart showing the bird strikes by phase of light.





```
altitude_effect = data.groupby('Altitude bin')['Effect: Impact to flight'].value_counts().unstack().fillna(0)
plt.figure(figsize=(12, 8))
altitude_effect.plot(kind='bar', stacked=True, colormap='viridis')
plt.title('Effect of Bird Strikes at Different Altitudes')
plt.xlabel('Altitude bin (feet)')
plt.ylabel('Number of Strikes')
plt.show()
```

## Effect of bird strikes on flight

A plot chart showing the Effect of bird strikes on flight.



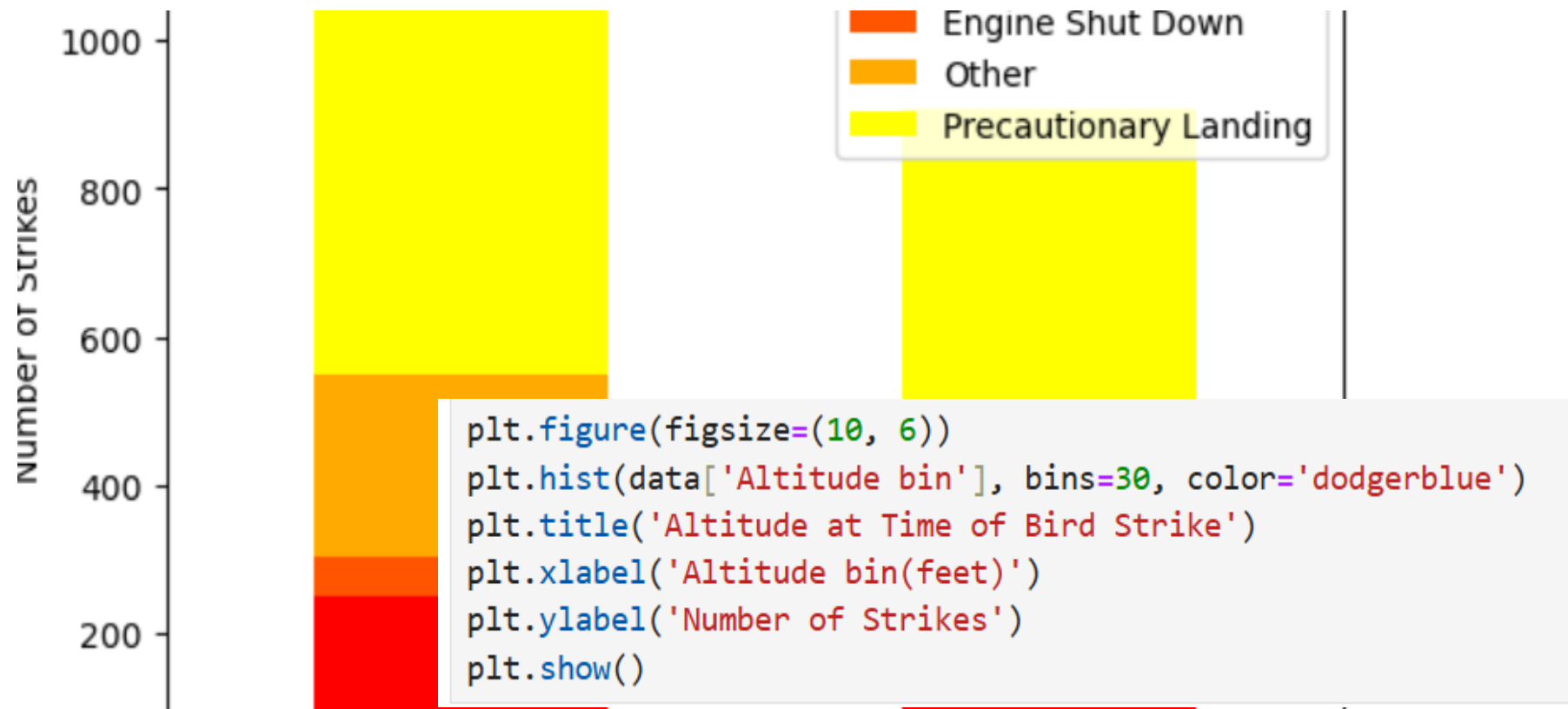
```
warning_effect = data.groupby('Pilot warned of birds or wildlife?')['Effect: Impact to flight'].value_counts().unstack().fillna(0)
plt.figure(figsize=(12, 8))
warning_effect.plot(kind='bar', stacked=True, colormap='autumn')
plt.title('Effect of Bird Strikes with Prior Warning')
plt.xlabel('Pilot warned of birds or wildlife?')
plt.ylabel('Number of Strikes')
plt.show()
```

<Figure size 1200x800 with 0 Axes>



## Effect of bird strikes at different altitudes

A stacked chart showing the effect of bird strikes at different altitudes.



## Effect of bird strikes with prior warning

A histogram chart showing the effect of bird strikes with prior warning.

The background features abstract, overlapping geometric shapes in various shades of blue, ranging from light sky blue to deep navy blue. These shapes are primarily located on the left and right sides of the frame, creating a modern, dynamic feel. The central area is a plain white background where the text is placed.

# **UNIFIED MENTOR INTERN PROJECT 2**

FIFA World Cup Analysis

```
world_cups = pd.read_csv(r'C:\Users\VIS\Downloads\FIFA WC data (1)\WorldCups.csv')
```

world\_cups

	Year	Country	Winner	Runners-Up	Third	Fourth	GoalsScored	QualifiedTeams	MatchesPlayed	Attendance
0	1930	Uruguay	Uruguay	Argentina	USA	Yugoslavia	70	13	18	590.549
1	1934	Italy	Italy	Czechoslovakia	Germany	Austria	70	16	17	363.000
2	1938	France	Italy	Hungary	Brazil	Sweden	84	15	18	375.700
3	1950	Brazil	Uruguay	Brazil	Sweden	Spain	88	13	22	1.045.246
4	1954	Switzerland	Germany FR	Hungary	Austria	Uruguay	140	16	26	768.607
5	1958	Sweden	Brazil	Sweden	France	Germany FR	126	16	35	819.810
6	1962	Chile	Brazil	Czechoslovakia	Chile	Yugoslavia	89	16	32	893.172
7	1966	England	England	Germany FR	Portugal	Soviet Union	89	16	32	1.563.135
8	1970	Mexico	Brazil	Italy	Germany FR	Uruguay	95	16	32	1.603.975
9	1974	Germany	Germany FR	Netherlands	Poland	Brazil	97	16	38	1.865.753

## World cup analysis data

A showing the FIFA world cup data to visualization the above the image.

```
ld_cup_matches = pd.read_csv(r'C:\Users\VIS\Downloads\FIFA WC data (1)\WorldCupMatches.csv')
ld_cup_matches
```

Year	Datetime	Stage	Stadium	City	Home Team Name	Home Team Goals	Away Team Goals	Away Team Name	Win conditions	...	Half-time Home Goals	Half-time Away Goals	Referee	Assistant 1	Assistant 2
1930	13 Jul 1930 - 15:00	Group 1	Pocitos	Montevideo	France	4	1	Mexico	...		3.0	0.0	LOMBARDI Domingo (URU)	CRISTOPHE Henry (BEL)	FRANK GILKES (ENG)
1930	13 Jul 1930 - 15:00	Group 4	Parque Central	Montevideo	USA	3	0	Belgium	...		2.0	0.0	MACIAS Jose (ARG)	MATEUCCI Francisco (URU)	WARNKEN Alberto (CHI)
1930	14 Jul 1930 - 12:45	Group 2	Parque Central	Montevideo	Yugoslavia	2	1	Brazil	...		2.0	0.0	TEJADA Anibal (URU)	VALLARINO Ricardo (URU)	BALBUENA Theodor (GER)
1930	14 Jul 1930 - 14:50	Group 3	Pocitos	Montevideo	Romania	3	1	Peru	...		1.0	0.0	WARNKEN Alberto (CHI)	LANGENUS Jean (BEL)	MATEUCCI Francisco (URU)
1930	15 Jul 1930 - 16:00	Group 1	Parque Central	Montevideo	Argentina	1	0	France	...		0.0	0.0	REGO Gilberto (BRA)	SAUCEDO Ulises (BOL)	RADULI Constantin (ROM)

## World cup matches

A showed the world cup matches the data visualization.

```

rld_cup_matches['Year'] = world_cup_matches['Year'].astype(int)
rld_cup_matches['Home Team Goals'] = world_cup_matches['Home Team Goals'].astype(int)
rld_cup_matches['Away Team Goals'] = world_cup_matches['Away Team Goals'].astype(int)

rld_cup_matches['Goal Difference'] = world_cup_matches['Home Team Goals'] - world_cup_matches['Away Team Goals']

rld_cups.to_csv(r'C:\Users\VIS\Downloads\FIFA WC data (1)\WorldCups.csv', index=False)
rld_cup_matches.to_csv(r'C:\Users\VIS\Downloads\FIFA WC data (1)\WorldCupMatches.csv', index=False)

am_stats = world_cup_matches.groupby(['Year', 'Home Team Name'])['Goal Difference'].mean().reset_index()

rged_data = pd.merge(team_stats, world_cups[['Year', 'Winner']], left_on=['Year', 'Home Team Name'], right_on=['Year', 'Winner'], how='left')

rged_data['Winner'] = merged_data['Winner'].notna().astype(int)

rged_data.to_csv('merged_data.csv', index=False)

p_scorers = world_cup_matches.groupby('Home Team Name')['Home Team Goals'].sum().reset_index()

p_scorers = top_scorers.sort_values(by='Home Team Goals', ascending=False).head(10)

p_scorers.to_csv('top_scorers.csv', index=False)

```

## Data format

A showing the above all the code of group by ,merge the data , and sort the values of the top scores values.

world\_cups\_players

	RoundID	MatchID	Team Initials	Coach Name	Line-up	Shirt Number	Player Name	Position	Event
0	201	1096	FRA	CAUDRON Raoul (FRA)	S	0	Alex THEPOT	GK	NaN
1	201	1096	MEX	LUQUE Juan (MEX)	S	0	Oscar BONFIGLIO	GK	NaN
2	201	1096	FRA	CAUDRON Raoul (FRA)	S	0	Marcel LANGILLER	NaN	G40'
3	201	1096	MEX	LUQUE Juan (MEX)	S	0	Juan CARRENO	NaN	G70'
4	201	1096	FRA	CAUDRON Raoul (FRA)	S	0	Ernest LIBERATI	NaN	NaN
...	...	...	...	...	...	...	...	...	...
37779	255959	300186501	ARG	SABELLA Alejandro (ARG)	N	19	ALVAREZ	NaN	NaN
37780	255959	300186501	GER	LOEW Joachim (GER)	N	6	KHEDIRA	NaN	NaN
37781	255959	300186501	ARG	SABELLA Alejandro (ARG)	N	20	AGUERO	NaN	IH46' Y65'
37782	255959	300186501	GER	LOEW Joachim (GER)	N	21	MUSTAFI	NaN	NaN
37783	255959	300186501	ARG	SABELLA Alejandro (ARG)	N	23	BASANTA	NaN	NaN

37784 rows x 9 columns

## World cups players data

A showed the above the world cup players of the data's.



)							
ry	Winner	Runners-Up	Third	Fourth	GoalsScored	QualifiedTeams	MatchesPlayed
ay	Uruguay	Argentina	USA	Yugoslavia	70	13	18
aly	Italy	Czechoslovakia	Germany	Austria	70	16	17
ice	Italy	Hungary	Brazil	Sweden	84	15	18
razil	Uruguay	Brazil	Sweden	Spain	88	13	22
nd	Germany FR	Hungary	Austria	Uruguay	140	16	26
s							

## World cup data

A showing world cup of head count.

```
Fourth          object
GoalsScored     int64
QualifiedTeams  int64
MatchesPlayed   int64
Attendance       object
dtype: object
```

```
world_cups.isnull().sum()
```

```
Year          0
Country       0
Winner        0
Runners-Up    0
```

## World cup data

A showing the is null value of the data and sum of the count.

```
Germany  1  
Uruguay  2  
Argentina 2  
England  1  
France   1  
Spain    1  
Germany  1  
Name: count, dtype: int64
```

```
runners_up = world_cups['Runners-Up'].value_counts()
```

```
runners_up
```

## World cup data

A showing the world cups data to show the Runner up data and the value of count.

<b>Brazil</b>	5	2	2
<b>Italy</b>	4	2	1
<b>Germany FR</b>	3	3	1
<b>Uruguay</b>	2	0	0
<b>Argentina</b>	2	3	0
<b>England</b>	1	0	0
<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>

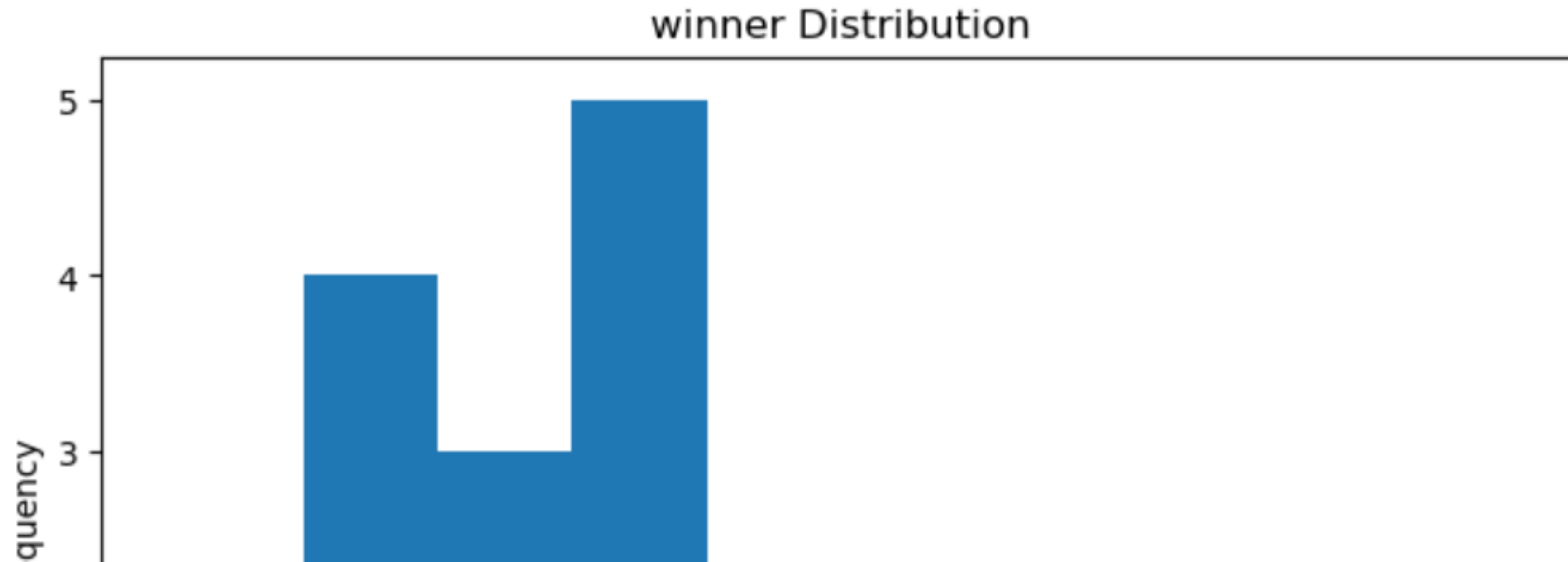
## World cup data

A showing the Teams of the data to display to the statement of below.

```
teams = pd.concat([Winner, runners_up, third], axis=1) teams.fillna(0, inplace=True) teams = teams.astype(int)
```

```
plt.title("winner Distribution")
plt.xlabel("Winner")
plt.ylabel("Frequency")
```

```
: Text(0, 0.5, 'Frequency')
```

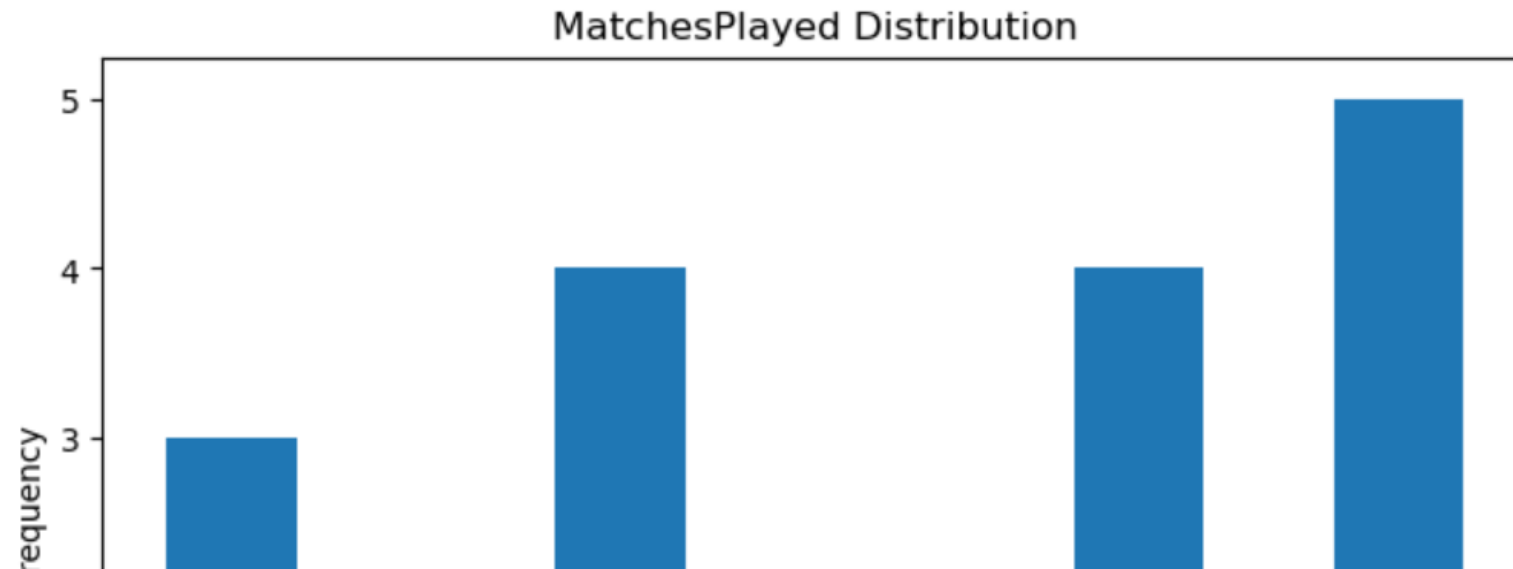


## World cup data

A showing the world cup data from matplotlib import pyplot as plt to visualization the winner of distribution.

```
plt.title("MatchesPlayed Distribution")  
plt.xlabel("MatchesPlayed")  
plt.ylabel("Frequency")
```

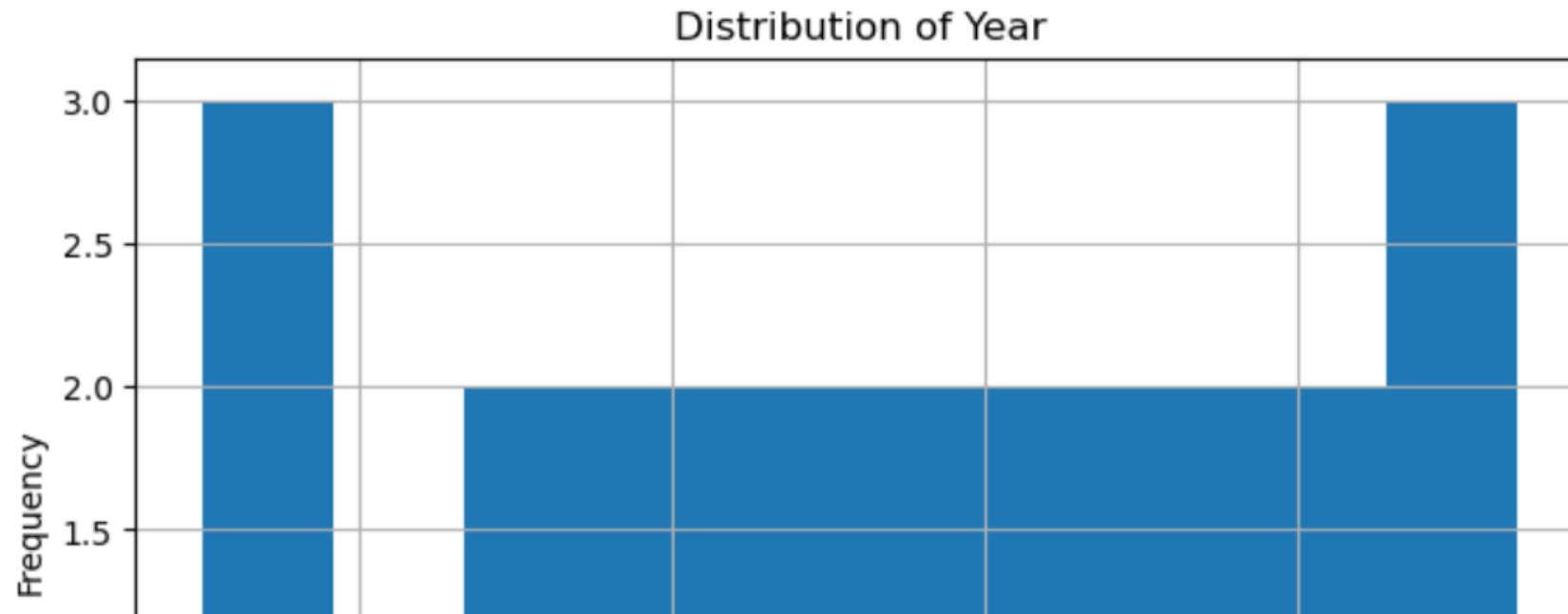
```
Text(0, 0.5, 'Frequency')
```



## World cup data

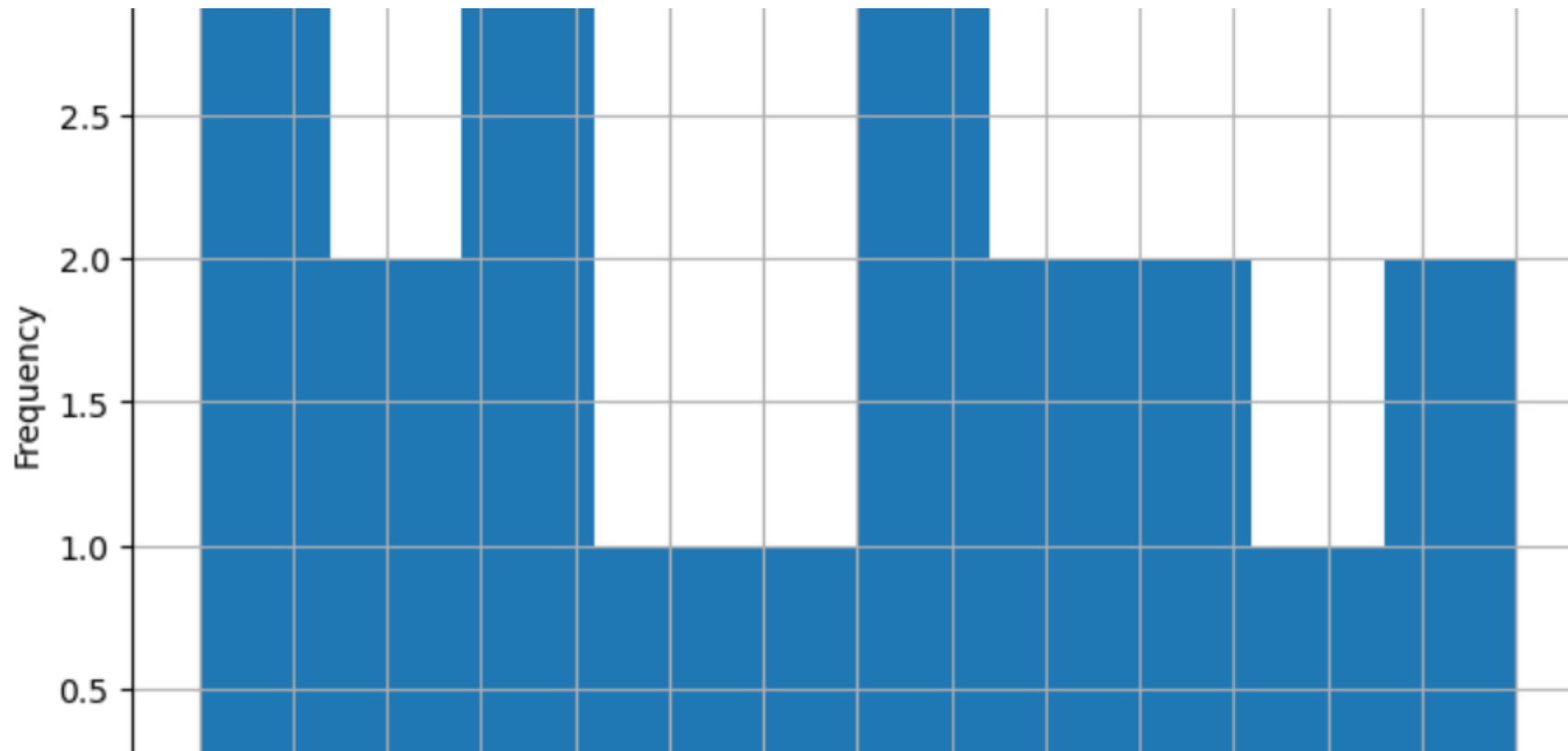
A showing the data like matches played distribution of the histogram chart.

```
plt.xlabel(column)
plt.ylabel("Frequency")
plt.show()
```



## World cup data

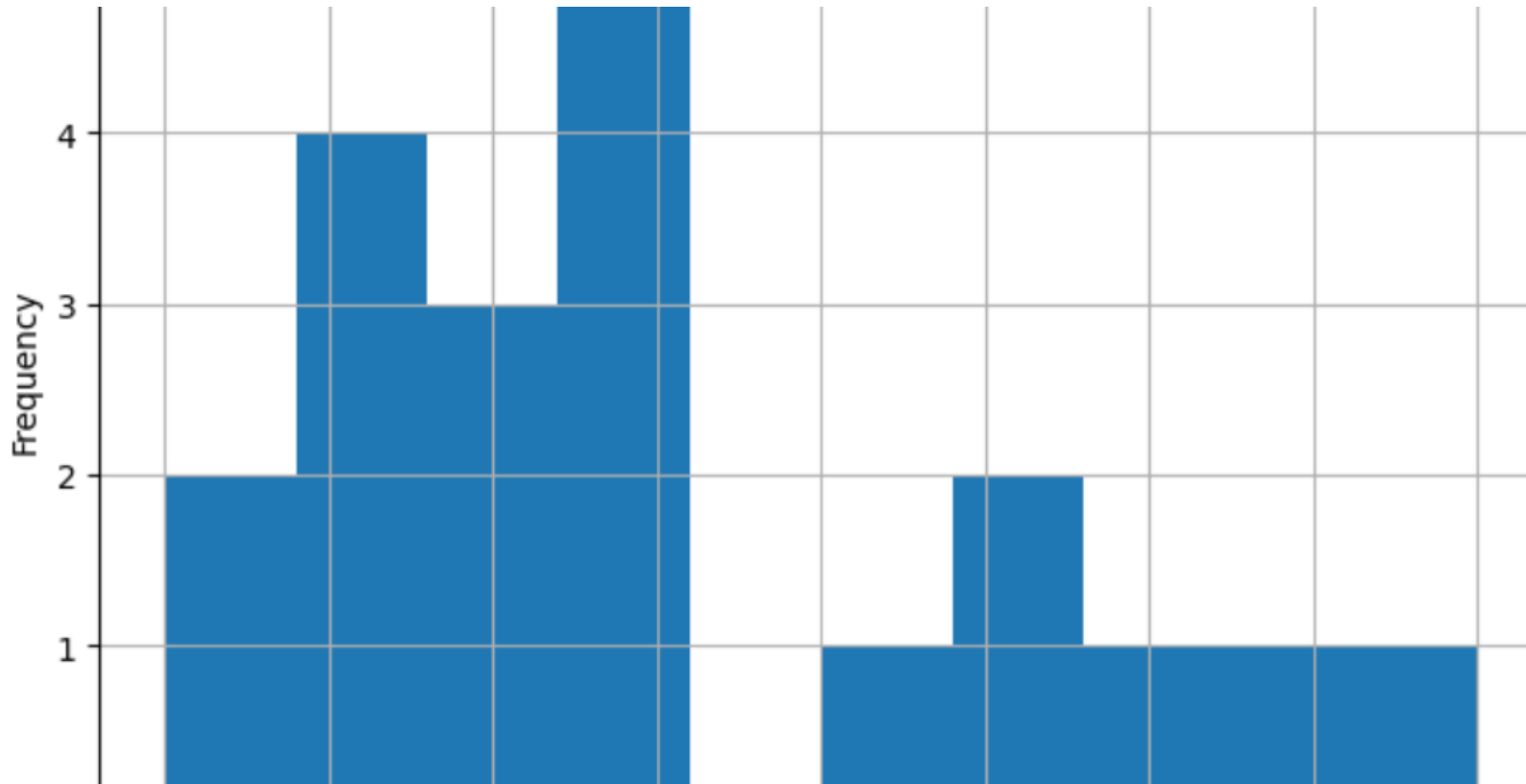
A showing the world cup data to visualization the histogram of distribution of year and frequency.



## World cup data

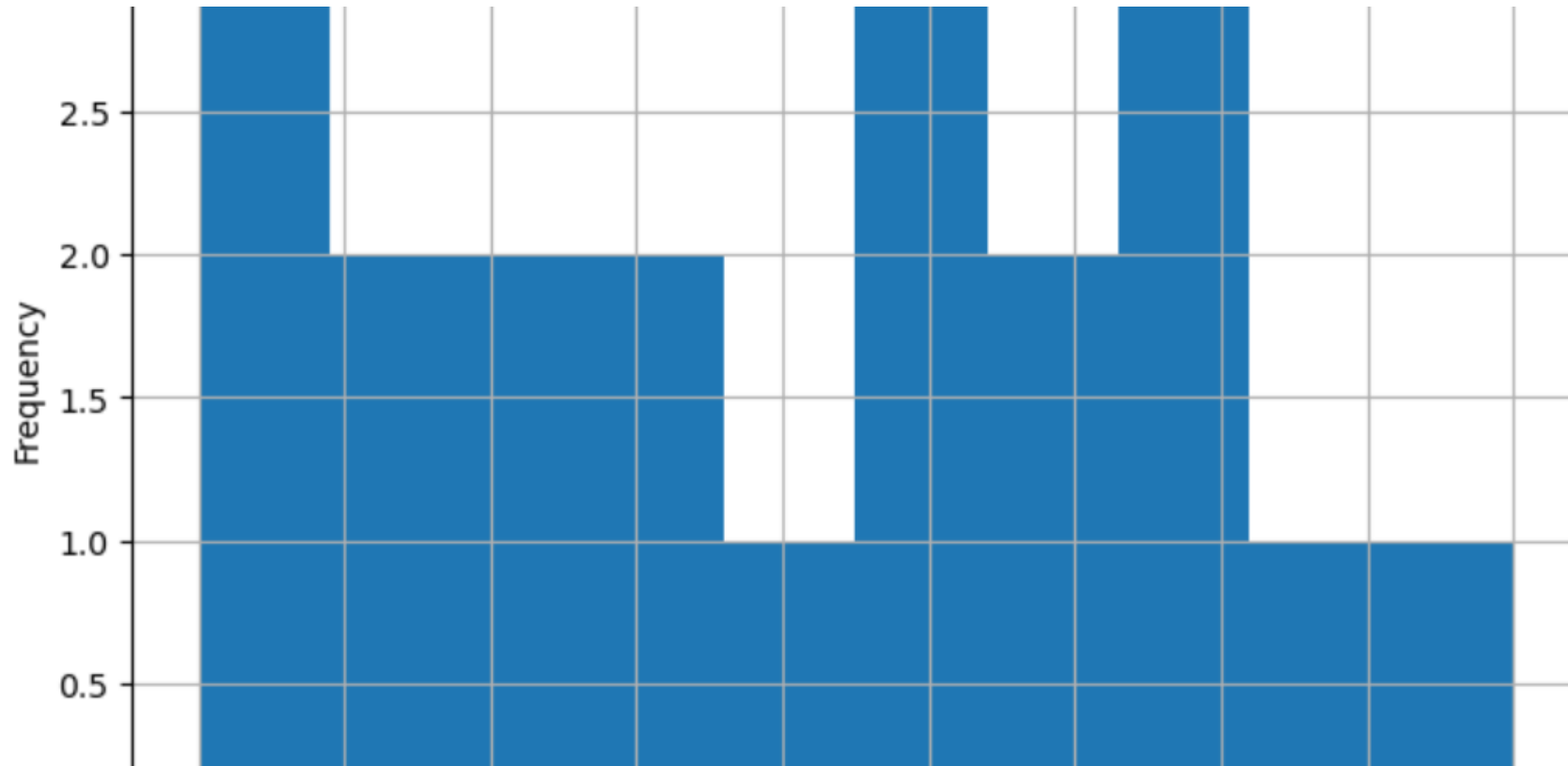
A showing the world cup data to visualization the histogram of distribution of country and frequency.





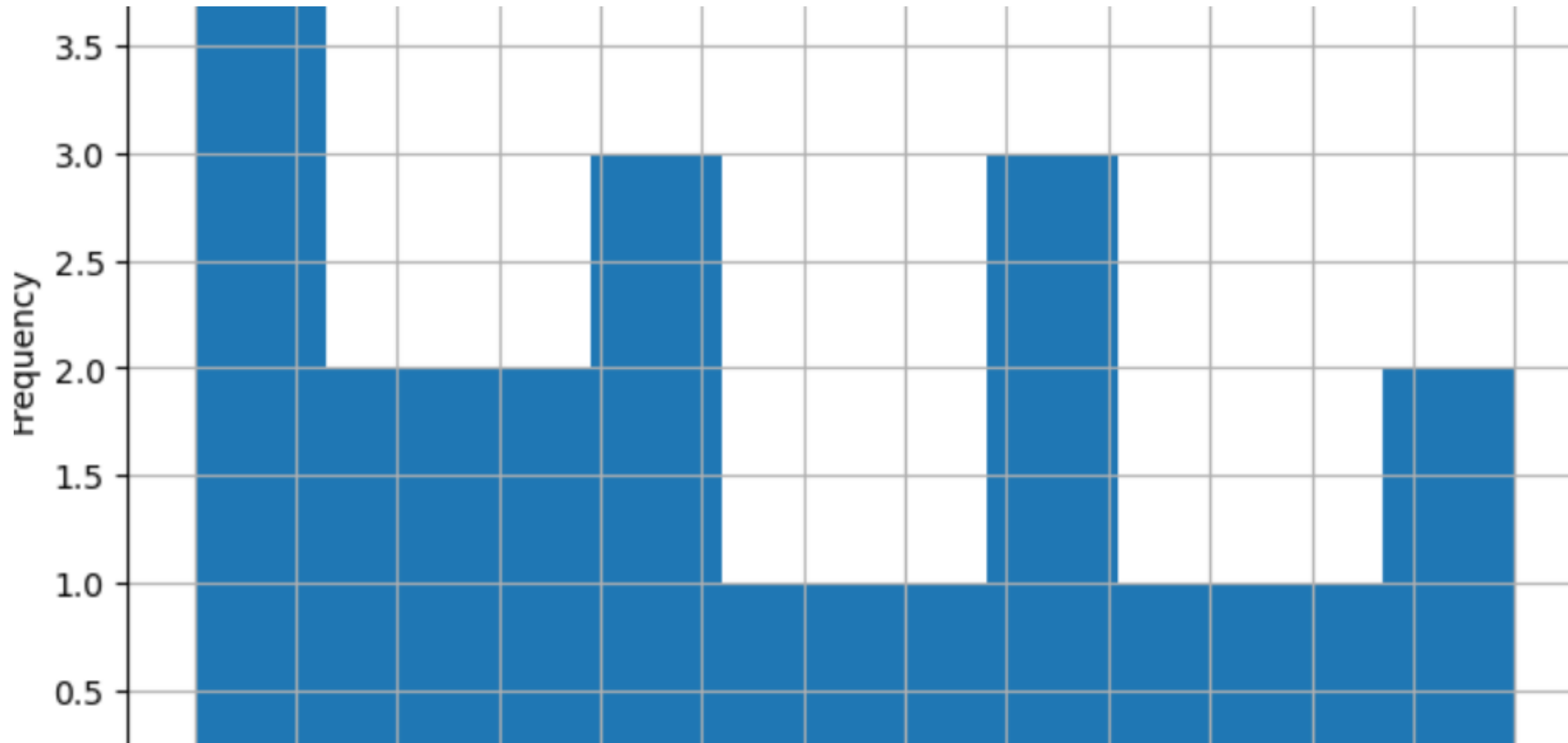
## World cup data

A showing the world cup data to visualization the histogram of distribution of winner and frequency.



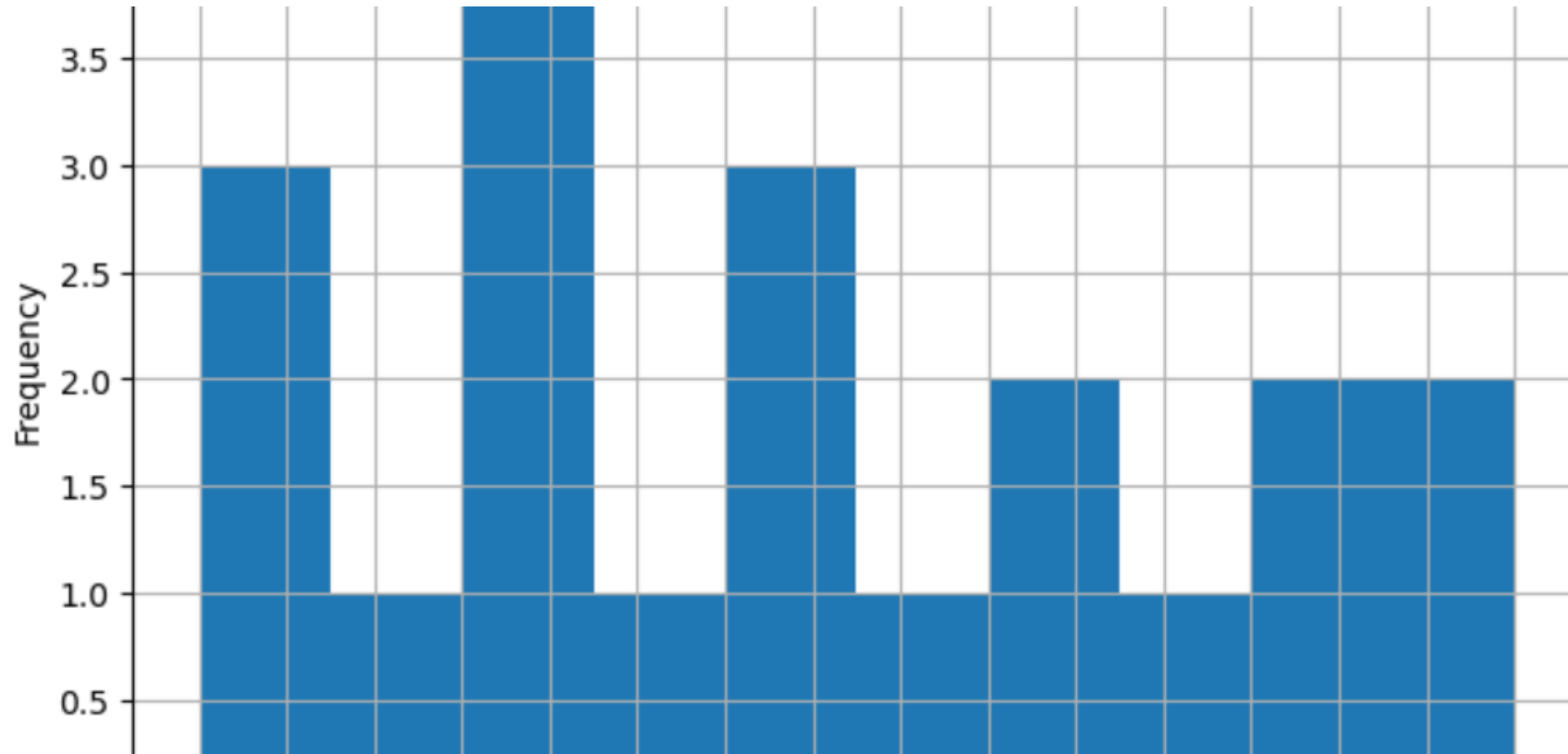
## World cup data

A showing the world cup data to visualization the histogram of distribution of Runner up and frequency.



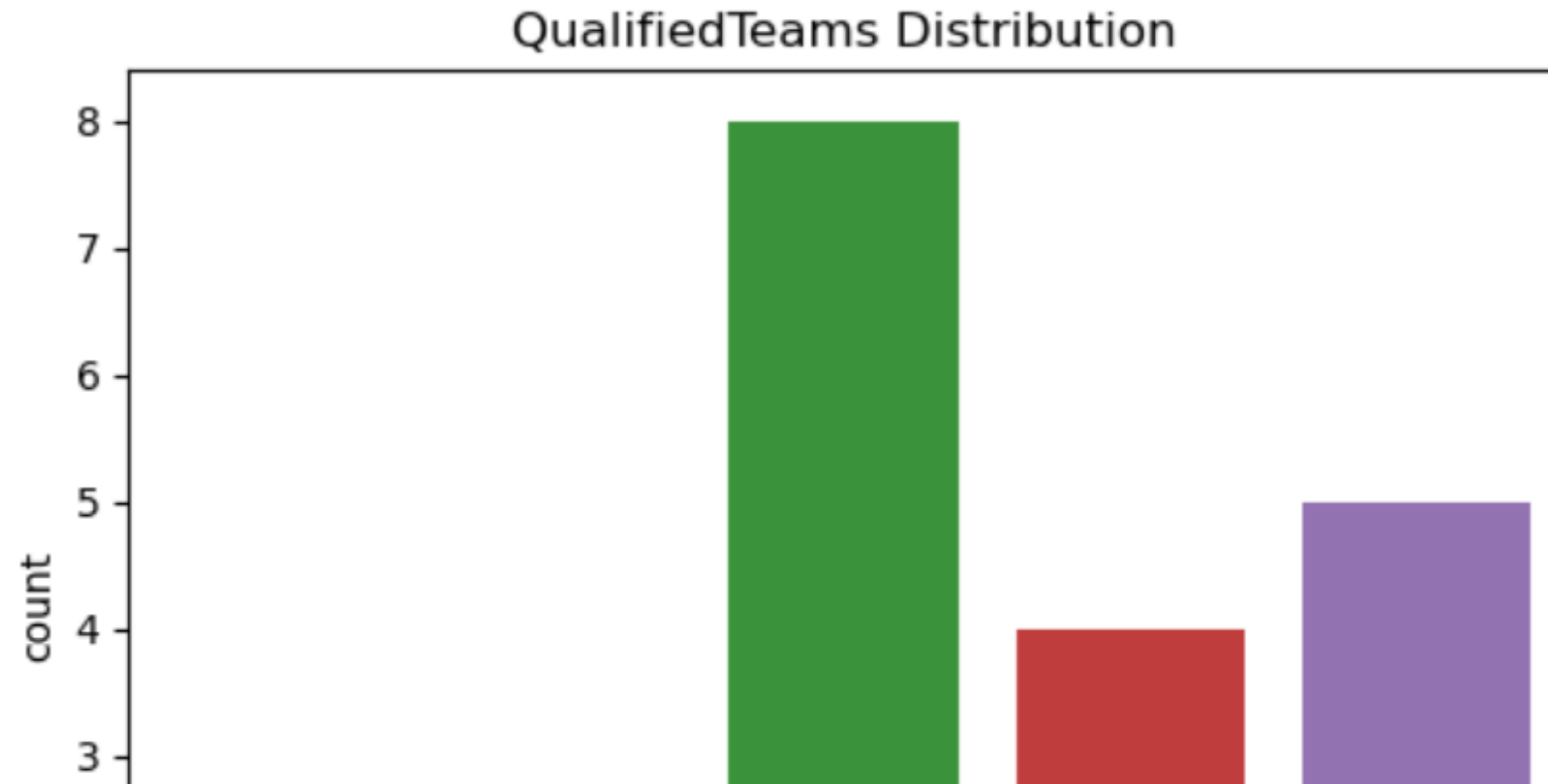
## World cup data

A showing the world cup data to visualization the histogram of distribution of Third and frequency.



## World cup data

A showing the world cup data to visualization the histogram of distribution of fourth and frequency.



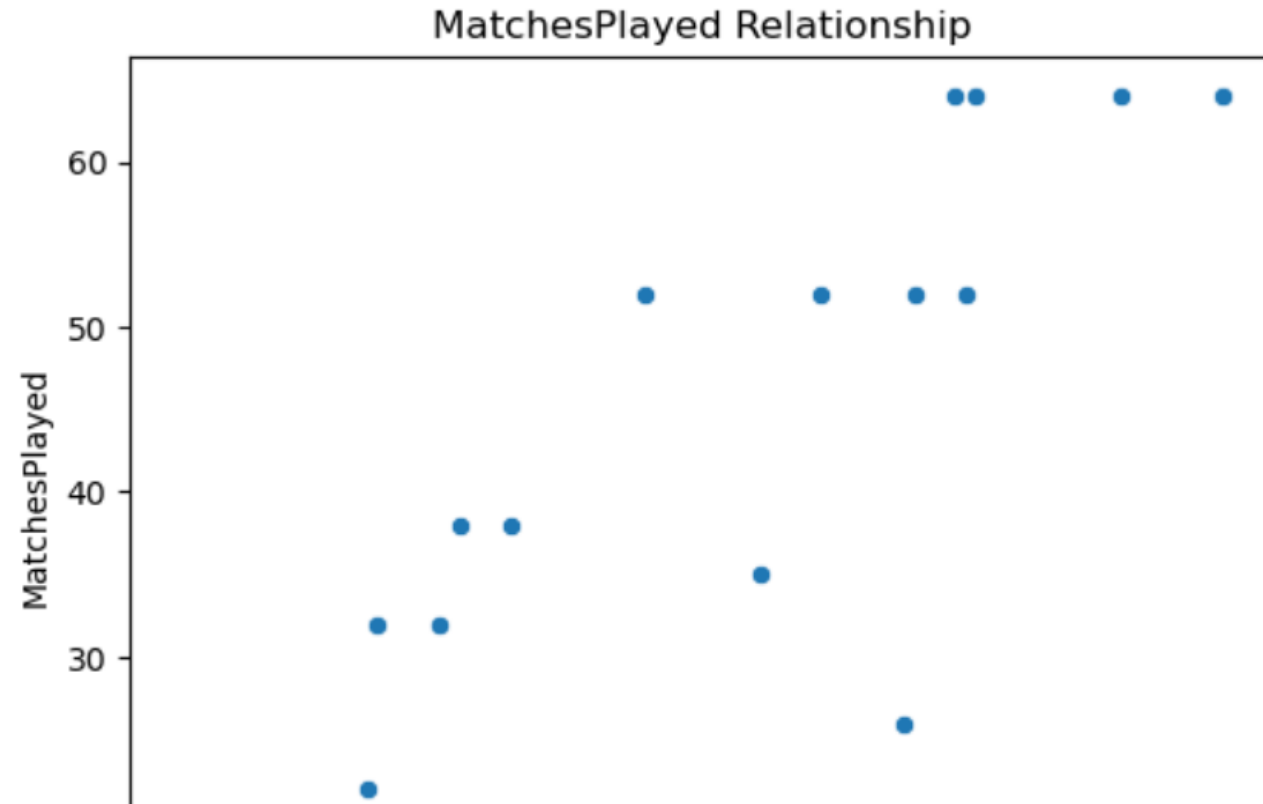
## World cup data

A showing the world cup data to visualization the countplot of distribution of Qualified teams and count.

<b>count</b>	20.000000	20.000000	20.000000	20.000000
<b>mean</b>	1974.800000	118.950000	21.250000	41.800000
<b>std</b>	25.582889	32.972836	7.268352	17.218717
<b>min</b>	1930.000000	70.000000	13.000000	17.000000
<b>25%</b>	1957.000000	89.000000	16.000000	30.500000
<b>50%</b>	1976.000000	120.500000	16.000000	38.000000
<b>75%</b>	1995.000000	145.250000	26.000000	55.000000
<b>max</b>	2014.000000	171.000000	32.000000	64.000000

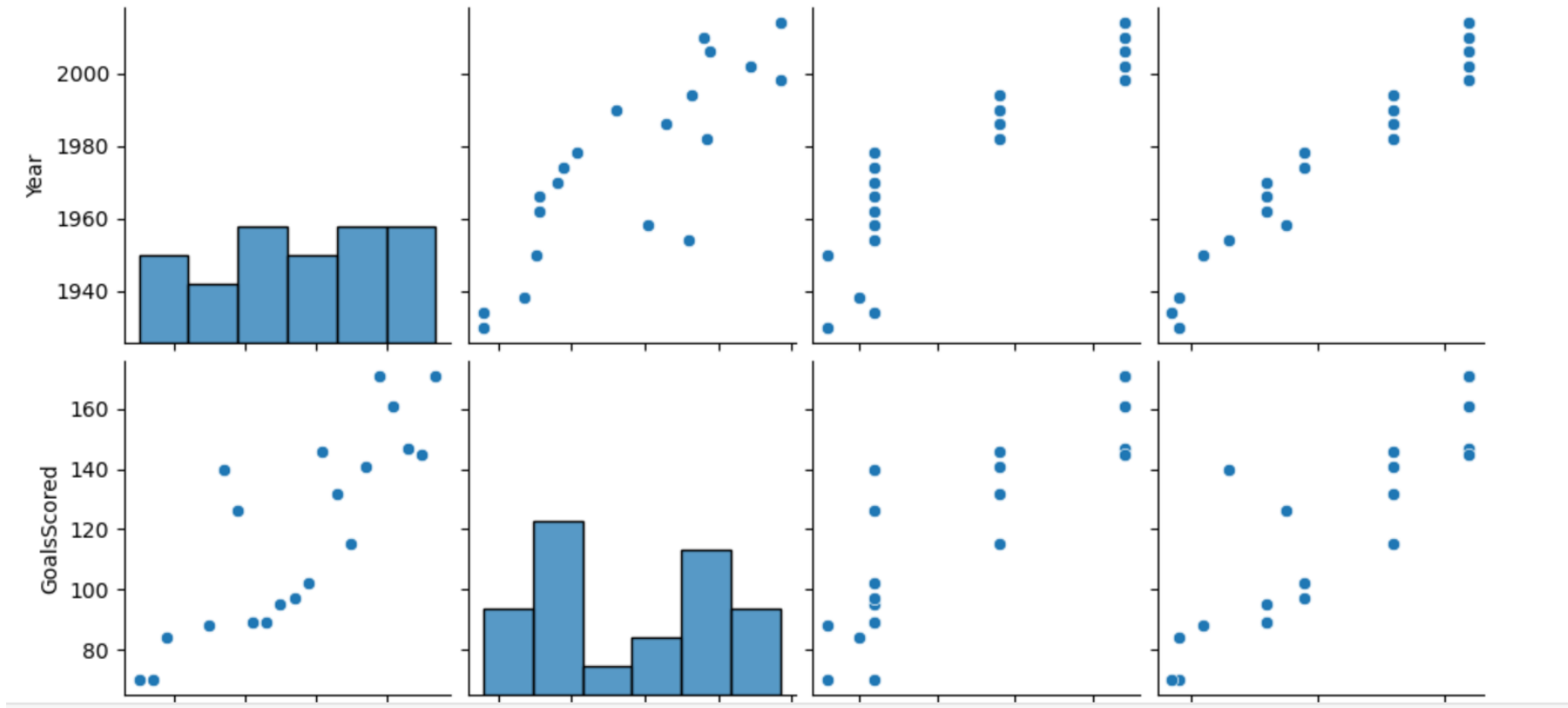
## World cup data

To describe the world cup data.



## World cup data

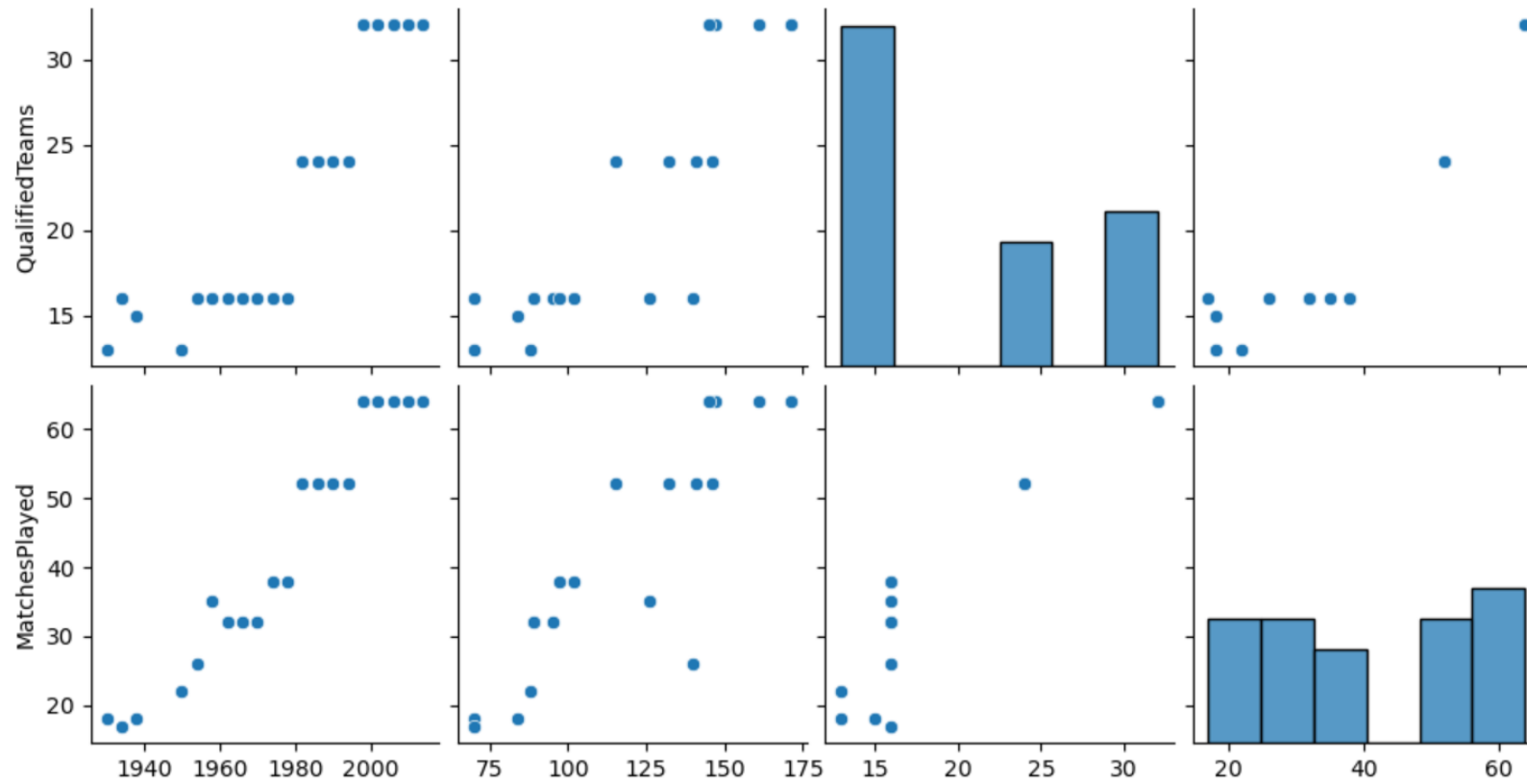
A showing the world cup data to visualization the Scatterplot of matches played relationship and matches played.



## World cup data

A showing the world cup data to visualization the pair plot of world cup data .





## World cup data

A showing the world cup data to visualization the Pair plot to the world cup data



**THANK YOU**