
 Marwadi University Marwadi Chandarana Group		Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Probability and Statistics (01CT1401)	Aim: Visualizations		
Task :- 2	Date:- 26-02-2024	Enrollment No:- 92200133030	

Day 2: Visualizations

Exercise – 1 :- Histogram

Code :-

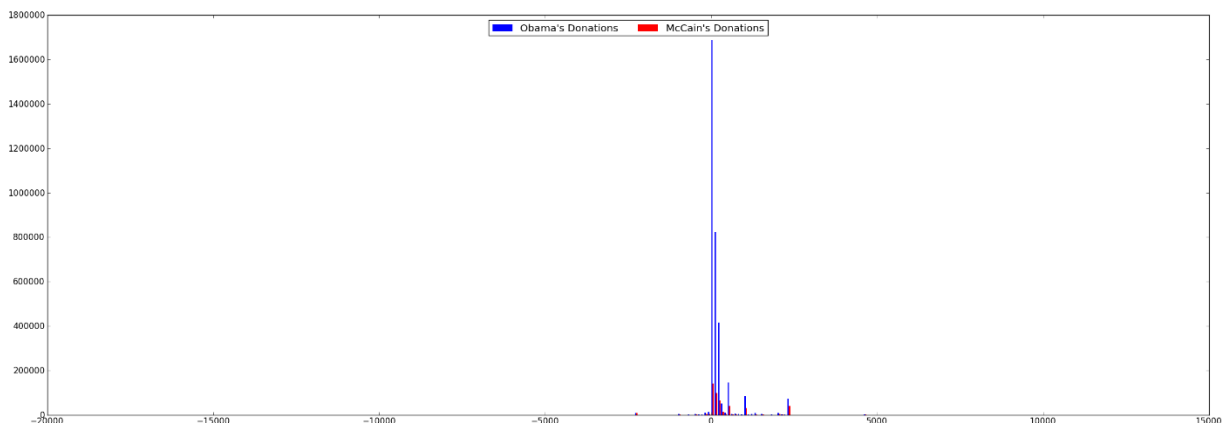
```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
data = pd.read_csv("./donations.csv", dtype={"contbr_zip": str, "memo_cd": str})
obama_data = data[data["cand_nm"] == "Obama, Barack"]
competitor_data = data[data["cand_nm"] == "McCain, John S."]
bins = np.arange(0, max(data["contb_receipt_amt"]) + 100, 100)
plt.figure(figsize=(10, 6))
plt.hist(obama_data["contb_receipt_amt"], bins=bins, alpha=0.5, label="Obama", color="blue")
plt.hist(competitor_data["contb_receipt_amt"], bins=bins, alpha=0.5, label="McClain", color="red")
mean_amt = data["contb_receipt_amt"].mean()
std_dev = data["contb_receipt_amt"].std()


plt.xlim(mean_amt - 3 * std_dev, mean_amt + 3 * std_dev)

plt.xlabel("Contribution Amount ($)")
plt.ylabel("Number of Contributions")
plt.title("Histogram of Contribution Amounts for Obama and Mc Clain")
plt.legend()

plt.grid(True)
plt.show()
```

Output :-



 Marwadi University Marwadi Chandarana Group	NAAC A+	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Probability and Statistics (01CT1401)	Aim: Visualizations		
Task :- 2	Date:- 26-02-2024	Enrollment No:- 92200133030	

Exercise – 1 :- Histogram

Code :-

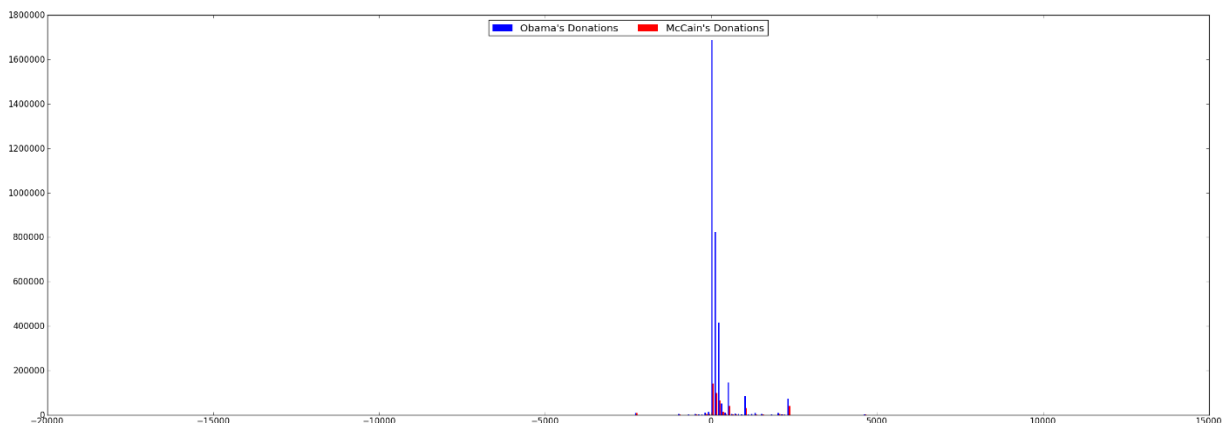
```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
data = pd.read_csv("./donations.csv", dtype={"contbr_zip": str, "memo_cd": str})
obama_data = data[data["cand_nm"] == "Obama, Barack"]
competitor_data = data[data["cand_nm"] == "McCain, John S."]
bins = np.arange(0, max(data["contb_receipt_amt"]) + 100, 100)
plt.figure(figsize=(10, 6))
plt.hist(obama_data["contb_receipt_amt"], bins=bins, alpha=0.5, label="Obama", color="blue")
plt.hist(competitor_data["contb_receipt_amt"], bins=bins, alpha=0.5, label="McClain", color="red")
mean_amt = data["contb_receipt_amt"].mean()
std_dev = data["contb_receipt_amt"].std()


plt.xlim(mean_amt - 3 * std_dev, mean_amt + 3 * std_dev)

plt.xlabel("Contribution Amount ($)")
plt.ylabel("Number of Contributions")
plt.title("Histogram of Contribution Amounts for Obama and Mc Clain")
plt.legend()

plt.grid(True)
plt.show()
```

Output :-



 Marwadi University Marwadi Chandarana Group	NAAC A+	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Probability and Statistics (01CT1401)	Aim: Visualizations		
Task :- 2	Date:- 26-02-2024	Enrollment No:- 92200133030	

Exercise – 2 :- More line graphs

Code :-

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

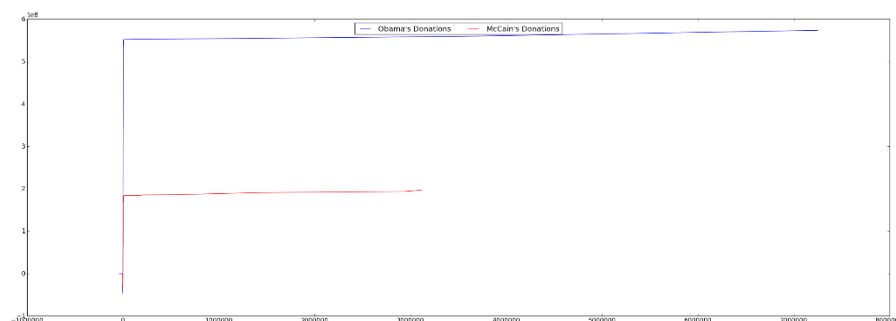
data = pd.read_csv("./donations.csv")
obama_data = data[data["cand_nm"] == "Obama, Barack"]
mccain_data = data[data["cand_nm"] == "McCain, John S."]


obama_data.sort_values("contb_receipt_amt", inplace=True)
mccain_data.sort_values("contb_receipt_amt", inplace=True)

obama_data["cumulative_donations"] = obama_data["contb_receipt_amt"].cumsum()
mccain_data["cumulative_donations"] = mccain_data["contb_receipt_amt"].cumsum()

plt.figure(figsize=(10, 6))
plt.plot(
    obama_data["contb_receipt_amt"], obama_data["cumulative_donations"], label="Obama"
)
plt.plot(
    mccain_data["contb_receipt_amt"],
    mccain_data["cumulative_donations"],
    label="McCain",
)
plt.xlabel("Donation Amount")
plt.ylabel("Cumulative Donations")
plt.title("Cumulative Donations for Obama and McCain")
plt.legend()
plt.grid(True)
plt.show()
```

Output :-



 Marwadi University Marwadi Chandarana Group	NAAC A+	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Probability and Statistics (01CT1401)	Aim: Visualizations		
Task :- 2	Date:- 26-02-2024	Enrollment No:- 92200133030	

Code :-

```

import pandas as pd
import matplotlib.pyplot as plt

data = pd.read_csv("./donations.csv")
obama_data = data[data["cand_nm"] == "Obama, Barack"]
mccain_data = data[data["cand_nm"] == "McCain, John S."]


obama_data.sort_values("contb_receipt_amt", inplace=True)
mccain_data.sort_values("contb_receipt_amt", inplace=True)

obama_data["cumulative_donations"] = obama_data["contb_receipt_amt"].cumsum()
mccain_data["cumulative_donations"] = mccain_data["contb_receipt_amt"].cumsum()

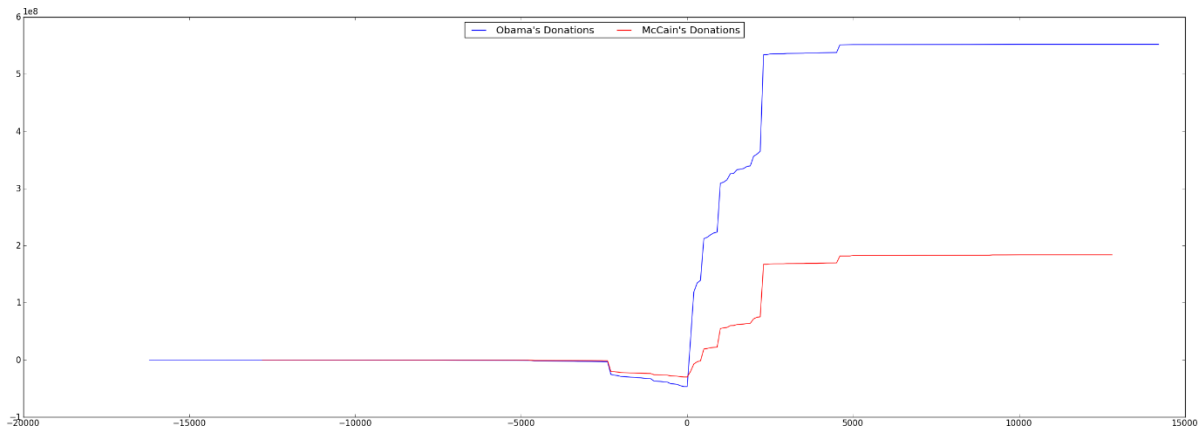
min_amount = min(obama_data["contb_receipt_amt"].min(), mccain_data["contb_receipt_amt"].min())
max_amount = max(obama_data["contb_receipt_amt"].max(), mccain_data["contb_receipt_amt"].max())
min_cumulative = min(obama_data["cumulative_donations"].min(), mccain_data["cumulative_donations"].min())
max_cumulative = max(obama_data["cumulative_donations"].max(), mccain_data["cumulative_donations"].max())

plt.figure(figsize=(10, 6))
plt.plot(
    obama_data["contb_receipt_amt"], obama_data["cumulative_donations"], label="Obama"
)
plt.plot(
    mccain_data["contb_receipt_amt"],
    mccain_data["cumulative_donations"],
    label="McCain",
)
plt.xlabel("Donation Amount")
plt.ylabel("Cumulative Donations")
plt.title("Cumulative Donations for Obama and McCain")
plt.legend()
plt.grid(True)
plt.xlim(min_amount, 1000)
plt.ylim(min_cumulative, max_cumulative)
plt.show()

```

 Marwadi University Marwadi Chandarana Group	NAAC A+	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Probability and Statistics (01CT1401)	Aim: Visualizations		
Task :- 2	Date:- 26-02-2024	Enrollment No:- 92200133030	

Output :-




Exercise – 3 :- More line graphs

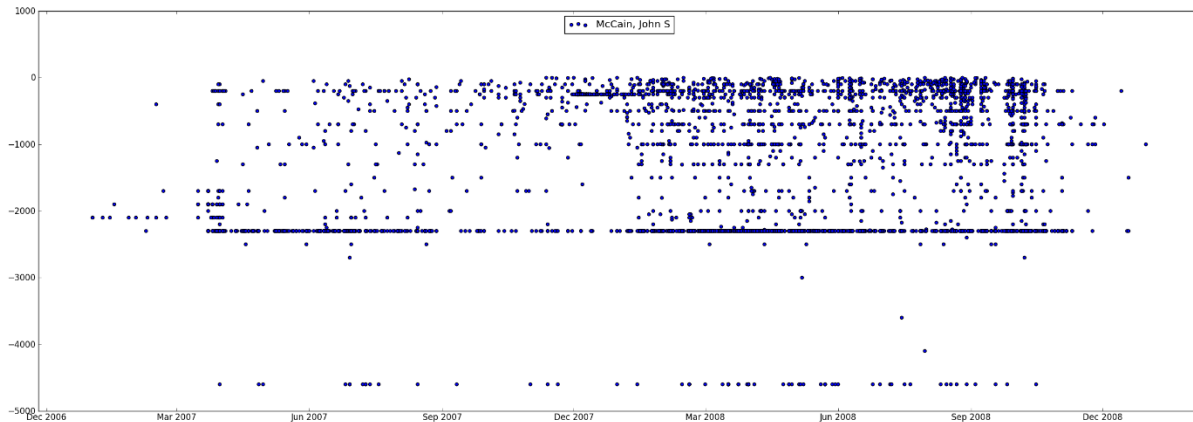
Code :-

```
import pandas as pd
import matplotlib.pyplot as plt
data = pd.read_csv("./donations.csv")
retribution_spouses = data[data['receipt_desc'] == 'REATTRIBUTION BY SPOUSE']
plt.figure(figsize=(10, 6))
for candidate in retribution_spouses['cand_nm'].unique():
    candidate_data = retribution_spouses[retribution_spouses['cand_nm'] == candidate]
    plt.scatter(candidate_data['contb_receipt_dt'], candidate_data['contb_receipt_amt'], label=candidate)

plt.xlabel('Donation Date')
plt.ylabel('Donation Amount')
plt.title('Re-attribution by Spouses for All Candidates')
plt.legend()
plt.grid(True)
plt.show()
```

 Marwadi University Marwadi Chandarana Group	NAAC A+	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Probability and Statistics (01CT1401)	Aim: Visualizations		
Task :- 2	Date:- 26-02-2024	Enrollment No:- 92200133030	

Output :-



Exercise – 4 :- More line graphs

Code :-

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

def get_color(donation_amount, max_donation):
    shade = 1 - (donation_amount / max_donation)
    return (0, 0, shade)

data = pd.read_csv("./donations.csv")


grouped = data.groupby(['cand_nm', 'contbr_st'])['contb_receipt_amt'].sum().reset_index()

candidates = grouped['cand_nm'].unique()

max_donation = grouped['contb_receipt_amt'].max()

num_candidates = len(candidates)
fig, axs = plt.subplots(num_candidates, figsize=(10, 6*num_candidates))

for i, candidate in enumerate(candidates):
    candidate_data = grouped[grouped['cand_nm'] == candidate]
    states = candidate_data['contbr_st']
```

 Marwadi University Marwadi Chandarana Group	NAAC A+	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Probability and Statistics (01CT1401)	Aim: Visualizations		
Task :- 2	Date:- 26-02-2024	Enrollment No:- 92200133030	

```
donation_amounts = candidate_data['contb_receipt_amt']
```

```
normalized_amounts = donation_amounts / max_donation
```

```
for state, amount in zip(states, normalized_amounts):
```

```
    color = get_color(amount, 1)
```

```
    axs[i].bar(state, 1, color=color, edgecolor='black')
```

```
axs[i].set_title(candidate)
```

```
axs[i].set_xlabel('State')
```

```
axs[i].set_ylabel('Total Donation')
```

```
axs[i].set_ylim(0, 1)
```

```
axs[i].set_yticks([])
```

```
plt.tight_layout()
```

```
plt.show()
```

Output :-

