CODE::

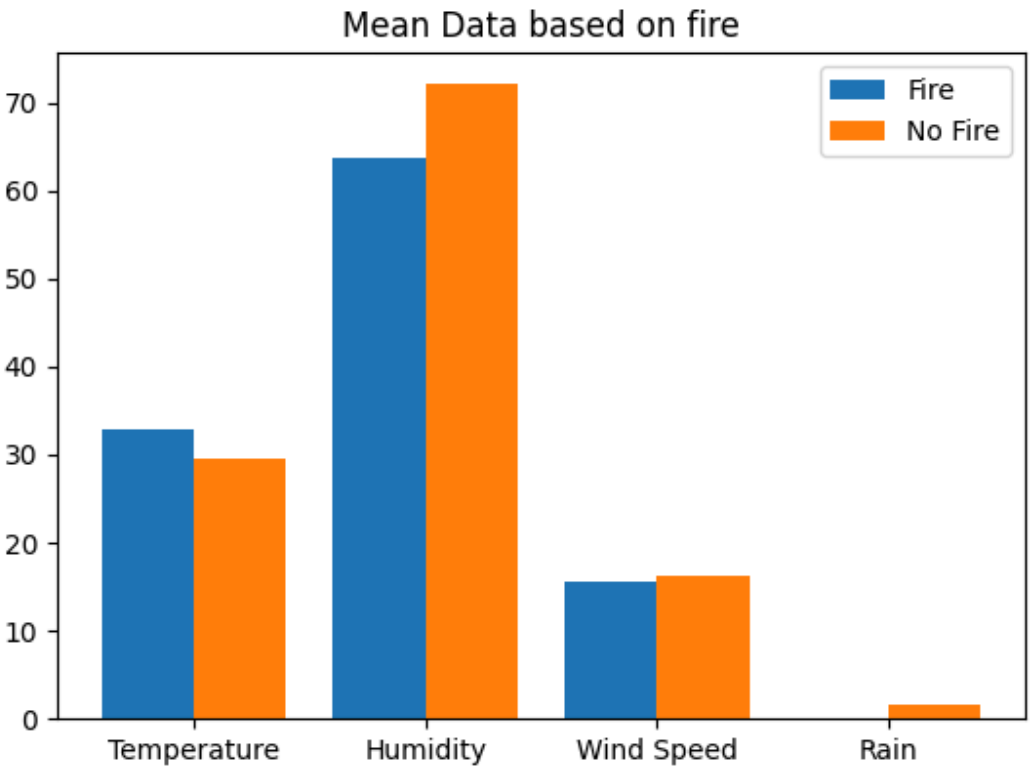
```
import stats as st
import pandas as pd
from matplotlib import pyplot as plt

sidiBelFrame = pd.read_csv("Sidi-Bel_Abbes_Region.csv")
bejaiaFrame = pd.read_csv("Bejaia_Region.csv")

#columnTitles = bejaia.columns


bejaia = bejaiaFrame.values.tolist()
sidiBel = sidiBelFrame.values.tolist()

for i in bejaia[:][:]:
        i[13] = i[13].strip()

#print(bejaia)

#for i in range(len(bejaia)):
#       print(bejaia[i][3])

bejFireTemps = [bejaia[i][3] for i in range(len(bejaia)) if bejaia[i][len(bejaia[i]) - 1] == 'fire']
bejFireHumid = [bejaia[i][4] for i in range(len(bejaia)) if bejaia[i][len(bejaia[i]) - 1] == 'fire']
bejFireWind = [bejaia[i][5] for i in range(len(bejaia)) if bejaia[i][len(bejaia[i]) - 1] == 'fire']
bejFireRain = [bejaia[i][6] for i in range(len(bejaia)) if bejaia[i][len(bejaia[i]) - 1] == 'fire']


bejNoFireTemps = [bejaia[i][3] for i in range(len(bejaia)) if bejaia[i][len(bejaia[i]) - 1] != 'fire']
bejNoFireHumid = [bejaia[i][4] for i in range(len(bejaia)) if bejaia[i][len(bejaia[i]) - 1] != 'fire']
bejNoFireWind = [bejaia[i][5] for i in range(len(bejaia)) if bejaia[i][len(bejaia[i]) - 1] != 'fire']
bejNoFireRain = [bejaia[i][6] for i in range(len(bejaia)) if bejaia[i][len(bejaia[i]) - 1] != 'fire']

#fireBejTemprs = [i for i in bejTemps if ]

width = .4
bejMeanCate = ["Temperature", "Humidity", "Wind Speed", "Rain"]
xArrangement = [i for i in range(len(bejMeanCate))]

plt.bar(xArrangement, [st.mean(bejFireTemps), st.mean(bejFireHumid), st.mean(bejFireWind),
st.mean(bejFireRain)], width, label='Fire')
plt.bar([i + width for i in xArrangement], [st.mean(bejNoFireTemps), st.mean(bejNoFireHumid),
st.mean(bejNoFireWind), st.mean(bejNoFireRain)], width, label='No Fire')
```

```
plt.xticks([i + width / 2 for i in xArrangement], bejMeanCate)

plt.title("Mean Data based on fire")

plt.legend()

plt.show()


'''

Observations:

        Temperatures increase During Fires
        Humidity Decreases During Fires
        Wind Speed might decrease during fires however its hard to tell given the small change
        Its hard to know the coorelation between rain and fires given the low amount of rain that
has found

'''


for i in sidiBel[:][:]:
        i[13] = i[13].strip()



sidiMedCate = ["FFMC Median", "DMC Median", "DC Median", "ISI Median"]

FFMCMed = st.median([sidiBel[i][7] for i in range(len(sidiBel))])
DMCMed = st.median([sidiBel[i][8] for i in range(len(sidiBel))])
DCMed = st.median([sidiBel[i][9] for i in range(len(sidiBel))])
ISIMed = st.median([sidiBel[i][10] for i in range(len(sidiBel))])

plt.bar(xArrangement, [FFMCMed, DMCMed, DCMed, ISIMed], width + .2, label = 'Sidi
Medians')

plt.xticks(xArrangement, sidiMedCate)
```

```
plt.show()


bejNoFireTemps = [bejaia[i][3] for i in range(len(bejaia)) if bejaia[i][len(bejaia[i]) - 1] != 'fire']

tempQuantiles = [
        st.quantile([bejaia[i][3] for i in range(len(bejaia))], .25),
        st.quantile([bejaia[i][3] for i in range(len(bejaia))], .6),
        st.quantile([bejaia[i][3] for i in range(len(bejaia))] , .75)
]

humidQuantiles = [
        st.quantile([bejaia[i][4] for i in range(len(bejaia))], .25),
        st.quantile([bejaia[i][4] for i in range(len(bejaia))], .6),
        st.quantile([bejaia[i][4] for i in range(len(bejaia))] , .75)
        ]

windQuantiles = [
        st.quantile([bejaia[i][5] for i in range(len(bejaia))], .25),
        st.quantile([bejaia[i][5] for i in range(len(bejaia))], .6),
        st.quantile([bejaia[i][5] for i in range(len(bejaia))] , .75)
        ]

rainQuantiles = [
        st.quantile([bejaia[i][6] for i in range(len(bejaia))], .25),
        st.quantile([bejaia[i][6] for i in range(len(bejaia))], .6),
        st.quantile([bejaia[i][6] for i in range(len(bejaia))] , .75)
        ]

print("Temperature Quantiles .25, .6, .75: ", tempQuantiles)
print("Relative Humidity Quantiles .25, .6, .75: ", humidQuantiles)
print("Wind Speed Quantiles .25, .6, .75: ", windQuantiles)
print("Rain Quantiles .25, .6, .75: ", rainQuantiles)


deviations = [
        st.std([sidiBel[i][3] for i in range(len(sidiBel))]),
        st.std([sidiBel[i][6] for i in range(len(sidiBel))]),
        st.std([sidiBel[i][11] for i in range(len(sidiBel))]),
        st.std([sidiBel[i][12] for i in range(len(sidiBel))]),
]
```

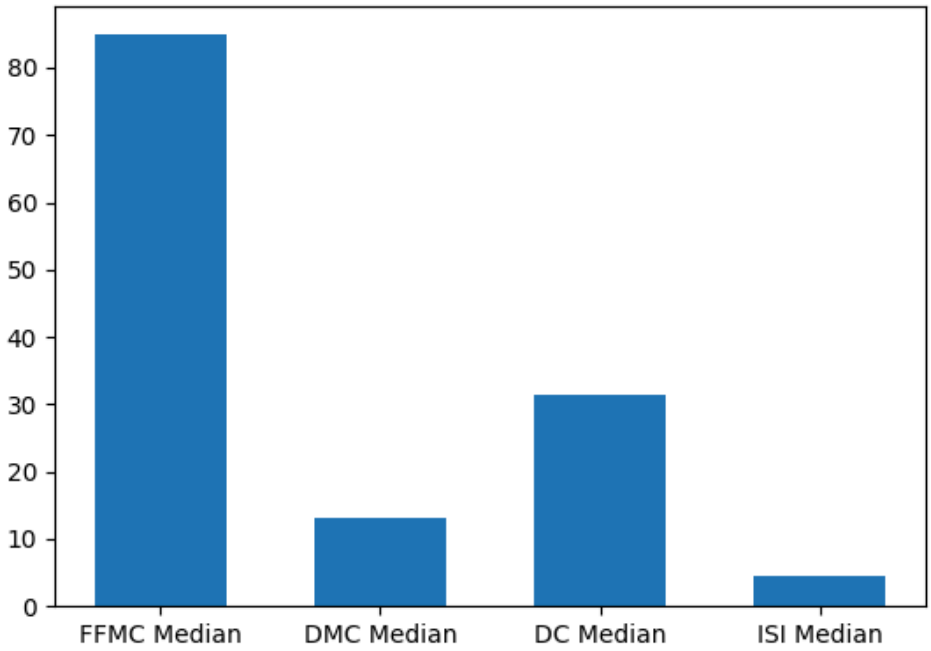
```
print("\nDeviations Temp, Rain, BUI, FWI: ", deviations)

RH = [bejaia[i][4] for i in range(len(sidiBel))]

coorelations = [
        st.correlation(RH, [bejaia[i][3] for i in range(len(bejaia))]),
        st.correlation(RH, [bejaia[i][5] for i in range(len(bejaia))]),
        st.correlation(RH, [bejaia[i][6] for i in range(len(bejaia))]),
        st.correlation(RH, [bejaia[i][7] for i in range(len(bejaia))]),
        st.correlation(RH, [bejaia[i][8] for i in range(len(bejaia))]),
        st.correlation(RH, [bejaia[i][9] for i in range(len(bejaia))]),
        st.correlation(RH, [bejaia[i][10] for i in range(len(bejaia))]),
        st.correlation(RH, [bejaia[i][11] for i in range(len(bejaia))]),
        st.correlation(RH, [bejaia[i][12] for i in range(len(bejaia))])
]

print("\nRelative Humidity compared to: Temperature, Wind Speed, Rain, FFMC, DMC, DC, ISI, BUI, FWI :", coorelations)



'''

        The strongest positive coorelation with Humidity is with Rain and the strongest negative coorelation is with Humidity and Temperature

'''



booleanFire = [int( bejaia[i][13] == 'fire') for i in range(len(bejaia))]

#dataCategories = [i for i in range(3, len(bejaia) - 1) if i != 2]

print("\nCoorelations:")
for i in range(3, len(bejaia[0]) - 1):
        print(i, st.correlation( booleanFire, [bejaia[j][i] for j in range(len(bejaia))]))
'''

        In this case I would choose to take the attributes FWI, ISI, FFMC
        This is because as computer with the code above the coorelation of these 3 data sets are the 3 highest in the dataset thus are the 3 best for differentiating

'''
```

Task 1:



From this graph we can see a positive correlation between Temperature and if there is a fire and a negative correlation between Humidity and fire. We also see their seems to be some negative correlation between fire and wind speed however it seems not as strong and could possibly just be some more static difference that is likely to exist we have a similar problem with rain especially since it seems there is a just a lot of days with little rain and it thus seems hard to tell if their just happened to be rain on fiery days or is their a correlation.

Task 2:



Task 3:

```
Temperature Quantiles .25, .6, .75:  [29, 32, 34]
Relative Humidity Quantiles .25, .6, .75:  [60, 73, 78]
Wind Speed Quantiles .25, .6, .75:  [14, 17, 18]
Rain Quantiles .25, .6, .75:  [0.0, 0.1, 0.5]
```

Task 4:

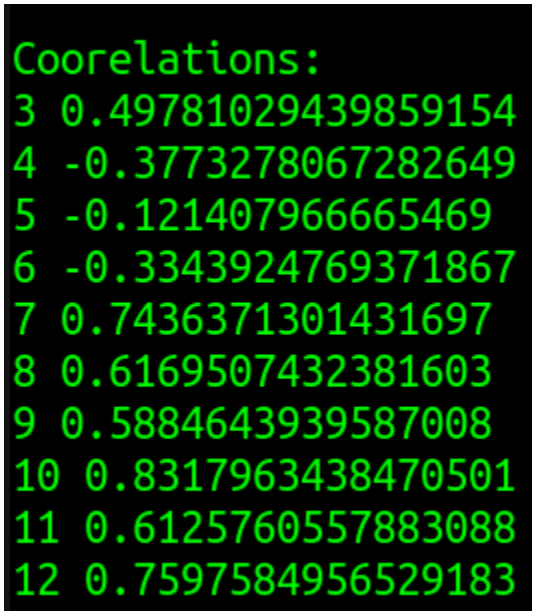
```
Deviations Temp, Rain, BUI, FWI:  [3.6605129438500184, 1.480652797336956, 13.81382022708875, 8.104005673421685]
```

Task 5:

```
Relative Humidity compared to: Temperature, Wind Speed, Rain, FFMC, DMC, DC, ISI, BUI, FWI : [-0.6601505504342499, 0.24577445895806885, 0.32916261703566
4, -0.6531529118901973, -0.34708013592219367, -0.3142712238307086, -0.5864099949298507, -0.33823263630375583, -0.4760673581458338]
```

We can see off of this that the strongest positive correlation was Humidities Correlation With rain and the Strongest negative correlation was with temperature although closely followed by FFMC.

Task 6:



This Data shows correlations between fire and each of the different attributes except year month and day as I didn't want to really consider the time but the conditions that went into a fire. We can see here that the highest correlations when you take the absolute value are: 10, 12, and 7. When looking at the data set this then correlates to the data attributes FWI, ISI, and FFMC respectively. Because of this high correlation it would be intuitive that these attributes would be the one I would look at in order to determine whether there is a fire or not. It is worth mentioning however that this doesn't take account for the fact that certain traits may have isolated correlation like for example perhaps the fires predicted by FWI and ISI are the same fires and thus making ISI less helpful of an data point so further exploration is available however these correlations are quite strong and I imagine we wouldn't really need to get any more previse than these 3.