Correlation Trials

Absenteeism at work

The dataset covers reasons for absence as well as various factors like day of week, seasons, transportation expense and more.

We will be looking at correlations between absences and these factors

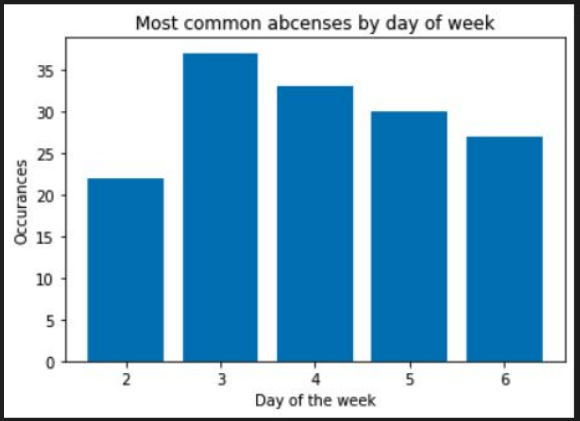
What is the most common reason for absence?

Overwhelmingly the most common reason for absence was reason 23, this chart displays the count of reason 23 by what day of the week it was on.

```
# sets x and y axis values to plot
x = absence_by_day['Day of the week']
y = absence_by_day['top absence Occurances']

plt.bar(x,y)

plt.title('Most common abcenses by day of week')

plt.xlabel('Day of the week')
plt.ylabel('Occurances')
```

✓ 0.1s

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



Why might people have this reason for absence?

The two most statistically significant reasons are both transport related.

```
spearmanr(df['Reason for absence'],
          df['Transportation expense'])
          #very little correlation, significant
```

More...

✓ 0.3s

```
SpearmanrResult(correlation=-0.12335601026796925, pvalue=0.0007714924987641616)
```

```
spearmanr(df['Reason for absence'],
          df['Distance from Residence to Work'])
          #very little correlation, significant
```

✓ 0.2s

```
SpearmanrResult(correlation=0.09913145161328486, pvalue=0.006959868996658 59)
```

How can we help prevent absences?

One way we can help with absences related to long transports is to incentivize getting to work early or on time.

Some ways to accomplish this may be:

-Free Breakfast that closes when it's time to work.

-Compensate employees for gas or other transport expenses. This will help alleviate any financial stress about getting to work.