

Lab 1

Objectives:

1. Introduce the concept of a data repository,
2. Introduce popular data repositories,
3. Extracting data from databases
4. Exploratory Data Analysis

Data Repositories

For your final project you need data. You can find good data sources online. The two most popular online sources are:

1. Data repositories: These are websites where people upload datasets for various purposes.
2. Data competition sites: Platforms where you can access datasets for data science competitions.

Data Source

Here are some websites you can collect your data from:

1. [The UC Irvine Machine Learning Repository](#)
2. [The Harvard Dataverse](#)
3. [FiveThirtyEight](#)
4. [NYtimes](#)
5. [Kaggle](#)
6. [The International Data Analysis Olympiad](#)

Check out these links for a few minutes!

Data Extraction

Let's see how you can use a repository to access data. Note that I'm not assuming you know any of these! So let me know if you have any questions.

Example 1

Check this out:

How Baby Boomers Get High

1. Go to the link associated with the data file
<https://github.com/fivethirtyeight/data/blob/master/drug-use-by-age/drug-use-by-age.csv>
2. Click on the **Raw** button above the data table on the page.
3. Save the file as `drug-use-by-age.csv`, and upload it to Google Colab
4. Run the code chunks below.

```
In [ ]: import pandas as pd
```

```
In [ ]: data = pd.read_csv("drug-use-by-age.csv")
```

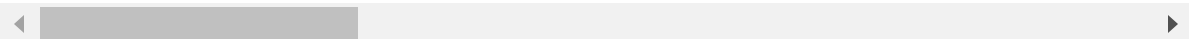
Let's look at the first 5 rows of our data

```
In [ ]: data.head()
```

```
Out[ ]:
```

	age	n	alcohol_use	alcohol_frequency	marijuana_use	marijuana_frequency	cocaine
0	12	2798	3.9	3.0	1.1	4.0	
1	13	2757	8.5	6.0	3.4	15.0	
2	14	2792	18.1	5.0	8.7	24.0	
3	15	2956	29.2	6.0	14.5	25.0	
4	16	3058	40.1	10.0	22.5	30.0	

5 rows × 28 columns



```
In [ ]: # Find the first 10 rows
# code here
print(data.head(10))
```

	age	n	alcohol_use	alcohol_frequency	marijuana_use	\
0	12	2798	3.9	3.0	1.1	
1	13	2757	8.5	6.0	3.4	
2	14	2792	18.1	5.0	8.7	
3	15	2956	29.2	6.0	14.5	
4	16	3058	40.1	10.0	22.5	
5	17	3038	49.3	13.0	28.0	
6	18	2469	58.7	24.0	33.7	
7	19	2223	64.6	36.0	33.4	
8	20	2271	69.7	48.0	34.0	
9	21	2354	83.2	52.0	33.0	

	marijuana_frequency	cocaine_use	cocaine_frequency	crack_use	\
0	4.0	0.1	5.0	0.0	
1	15.0	0.1	1.0	0.0	
2	24.0	0.1	5.5	0.0	
3	25.0	0.5	4.0	0.1	
4	30.0	1.0	7.0	0.0	
5	36.0	2.0	5.0	0.1	
6	52.0	3.2	5.0	0.4	
7	60.0	4.1	5.5	0.5	
8	60.0	4.9	8.0	0.6	
9	52.0	4.8	5.0	0.5	

	crack_frequency	...	oxycontin_use	oxycontin_frequency	tranquilizer_use	\
0	-	...	0.1	24.5	0.2	
1	3.0	...	0.1	41.0	0.3	
2	-	...	0.4	4.5	0.9	
3	9.5	...	0.8	3.0	2.0	
4	1.0	...	1.1	4.0	2.4	
5	21.0	...	1.4	6.0	3.5	
6	10.0	...	1.7	7.0	4.9	
7	2.0	...	1.5	7.5	4.2	
8	5.0	...	1.7	12.0	5.4	
9	17.0	...	1.3	13.5	3.9	

	tranquilizer_frequency	stimulant_use	stimulant_frequency	meth_use	\
0	52.0	0.2	2.0	0.0	
1	25.5	0.3	4.0	0.1	
2	5.0	0.8	12.0	0.1	
3	4.5	1.5	6.0	0.3	
4	11.0	1.8	9.5	0.3	
5	7.0	2.8	9.0	0.6	
6	12.0	3.0	8.0	0.5	
7	4.5	3.3	6.0	0.4	
8	10.0	4.0	12.0	0.9	
9	7.0	4.1	10.0	0.6	

	meth_frequency	sedative_use	sedative_frequency
0	-	0.2	13.0
1	5.0	0.1	19.0
2	24.0	0.2	16.5
3	10.5	0.4	30.0
4	36.0	0.2	3.0
5	48.0	0.5	6.5
6	12.0	0.4	10.0

7	105.0	0.3	6.0
8	12.0	0.5	4.0
9	2.0	0.3	9.0

[10 rows x 28 columns]

```
In [ ]: # Print the last 5 row?
# code here
print(data.tail(5))
```

	age	n	alcohol_use	alcohol_frequency	marijuana_use	\
12	26-29	2628	80.7	52.0	20.8	
13	30-34	2864	77.5	52.0	16.4	
14	35-49	7391	75.0	52.0	10.4	
15	50-64	3923	67.2	52.0	7.3	
16	65+	2448	49.3	52.0	1.2	

	marijuana_frequency	cocaine_use	cocaine_frequency	crack_use	\
12	52.0	3.2	5.0	0.4	
13	72.0	2.1	8.0	0.5	
14	48.0	1.5	15.0	0.5	
15	52.0	0.9	36.0	0.4	
16	36.0	0.0	-	0.0	

	crack_frequency	...	oxycontin_use	oxycontin_frequency	tranquilizer_use	\
12	6.0	...	1.2	13.5	4.2	
13	15.0	...	0.9	46.0	3.6	
14	48.0	...	0.3	12.0	1.9	
15	62.0	...	0.4	5.0	1.4	
16	-	...	0.0	-	0.2	

	tranquilizer_frequency	stimulant_use	stimulant_frequency	meth_use	\
12	10.0	2.3	7.0	0.6	
13	8.0	1.4	12.0	0.4	
14	6.0	0.6	24.0	0.2	
15	10.0	0.3	24.0	0.2	
16	5.0	0.0	364.0	0.0	

	meth_frequency	sedative_use	sedative_frequency
12	30.0	0.4	4.0
13	54.0	0.4	10.0
14	104.0	0.3	10.0
15	30.0	0.2	104.0
16	-	0.0	15.0

[5 rows x 28 columns]

```
In [ ]: #check what the function data.sample()does. What happens if you toss a number inside
print(data.sample())

"""
If you toss a number inside the parentheses, you will get a number of random entries
"""
```

```

    age      n  alcohol_use  alcohol_frequency  marijuana_use  \
11  24-25  4591           83.1             52.0           24.9

    marijuana_frequency  cocaine_use  cocaine_frequency  crack_use  \
11                   60.0           4.0             6.0           0.5

    crack_frequency  ...  oxycontin_use  oxycontin_frequency  tranquilizer_use  \
11                6.0  ...             1.3             20.0           4.3

    tranquilizer_frequency  stimulant_use  stimulant_frequency  meth_use  \
11                   10.0           2.6             10.0           0.7

    meth_frequency  sedative_use  sedative_frequency
11              21.0           0.2             17.5

```

[1 rows x 28 columns]

Lets have a overall look of the dataset, checking if there is any inconsistency in the Dataset.

```
In [ ]: print('Rows:',data.shape[0])
        print('Columns:',data.shape[1])
```

Rows: 17
Columns: 28

```
In [ ]: print('Columns (features)')
        print()
        print( data.columns.tolist())
```

Columns (features)

```

['age', 'n', 'alcohol_use', 'alcohol_frequency', 'marijuana_use', 'marijuana_frequen
cy', 'cocaine_use', 'cocaine_frequency', 'crack_use', 'crack_frequency', 'heroin_us
e', 'heroin_frequency', 'hallucinogen_use', 'hallucinogen_frequency', 'inhalant_us
e', 'inhalant_frequency', 'pain_releiver_use', 'pain_releiver_frequency', 'oxycontin
_use', 'oxycontin_frequency', 'tranquilizer_use', 'tranquilizer_frequency', 'stimula
nt_use', 'stimulant_frequency', 'meth_use', 'meth_frequency', 'sedative_use', 'sedat
ive_frequency']

```

```
In [ ]: print(' Unique Values')
        print()
        print( data.nunique())
```

Unique Values

age	17
n	17
alcohol_use	16
alcohol_frequency	9
marijuana_use	17
marijuana_frequency	10
cocaine_use	14
cocaine_frequency	10
crack_use	5
crack_frequency	13
heroin_use	9
heroin_frequency	17
hallucinogen_use	16
hallucinogen_frequency	6
inhalant_use	15
inhalant_frequency	11
pain_releiver_use	15
pain_releiver_frequency	10
oxycontin_use	12
oxycontin_frequency	15
tranquilizer_use	15
tranquilizer_frequency	10
stimulant_use	16
stimulant_frequency	11
meth_use	9
meth_frequency	14
sedative_use	6
sedative_frequency	14
dtype:	int64

Here is another useful function:

```
In [ ]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 17 entries, 0 to 16
Data columns (total 28 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   age                                   17 non-null     object
1   n                                     17 non-null     int64
2   alcohol_use                          17 non-null     float64
3   alcohol_frequency                    17 non-null     float64
4   marijuana_use                        17 non-null     float64
5   marijuana_frequency                  17 non-null     float64
6   cocaine_use                          17 non-null     float64
7   cocaine_frequency                    17 non-null     object
8   crack_use                            17 non-null     float64
9   crack_frequency                      17 non-null     object
10  heroin_use                            17 non-null     float64
11  heroin_frequency                      17 non-null     object
12  hallucinogen_use                     17 non-null     float64
13  hallucinogen_frequency                17 non-null     float64
14  inhalant_use                         17 non-null     float64
15  inhalant_frequency                   17 non-null     object
16  pain_releiver_use                    17 non-null     float64
17  pain_releiver_frequency               17 non-null     float64
18  oxycontin_use                        17 non-null     float64
19  oxycontin_frequency                  17 non-null     object
20  tranquilizer_use                     17 non-null     float64
21  tranquilizer_frequency                17 non-null     float64
22  stimulant_use                        17 non-null     float64
23  stimulant_frequency                  17 non-null     float64
24  meth_use                             17 non-null     float64
25  meth_frequency                       17 non-null     object
26  sedative_use                         17 non-null     float64
27  sedative_frequency                   17 non-null     float64
dtypes: float64(20), int64(1), object(7)
memory usage: 3.8+ KB
```

We can also load the directly from the website by placing the raw csv file directly into

```
.read_csv .
```

```
In [ ]: #Code here
data = pd.read_csv("https://raw.githubusercontent.com/fivethirtyeight/data/master/d
```

Can you filter your table to have information only among people in the age range 20-50

```
In [ ]: #Code here
import numpy as np
# print(data.loc['age'])
index_list = np.array([9,15])
print(data.iloc[9:15])
```

	age	n	alcohol_use	alcohol_frequency	marijuana_use	\
9	21	2354	83.2	52.0	33.0	
10	22-23	4707	84.2	52.0	28.4	
11	24-25	4591	83.1	52.0	24.9	
12	26-29	2628	80.7	52.0	20.8	
13	30-34	2864	77.5	52.0	16.4	
14	35-49	7391	75.0	52.0	10.4	

	marijuana_frequency	cocaine_use	cocaine_frequency	crack_use	\
9	52.0	4.8	5.0	0.5	
10	52.0	4.5	5.0	0.5	
11	60.0	4.0	6.0	0.5	
12	52.0	3.2	5.0	0.4	
13	72.0	2.1	8.0	0.5	
14	48.0	1.5	15.0	0.5	

	crack_frequency	...	oxycontin_use	oxycontin_frequency	tranquilizer_use	\
9	17.0	...	1.3	13.5	3.9	
10	5.0	...	1.7	17.5	4.4	
11	6.0	...	1.3	20.0	4.3	
12	6.0	...	1.2	13.5	4.2	
13	15.0	...	0.9	46.0	3.6	
14	48.0	...	0.3	12.0	1.9	

	tranquilizer_frequency	stimulant_use	stimulant_frequency	meth_use	\
9	7.0	4.1	10.0	0.6	
10	12.0	3.6	10.0	0.6	
11	10.0	2.6	10.0	0.7	
12	10.0	2.3	7.0	0.6	
13	8.0	1.4	12.0	0.4	
14	6.0	0.6	24.0	0.2	

	meth_frequency	sedative_use	sedative_frequency
9	2.0	0.3	9.0
10	46.0	0.2	52.0
11	21.0	0.2	17.5
12	30.0	0.4	4.0
13	54.0	0.4	10.0
14	104.0	0.3	10.0

[6 rows x 28 columns]

Let's plot this new table using scatter matrix.

```
In [ ]: #Code here
pd.plotting.scatter_matrix(data.iloc[9:15], figsize=(16,9), alpha=0.2)
```



```
c:\Users\death\anaconda3\envs\openmm\lib\site-packages\pandas\plotting\_matplotlib\m
isc.py:100: UserWarning: Attempting to set identical left == right == 52.0 results i
n singular transformations; automatically expanding.
    ax.set_xlim(boundaries_list[j])
c:\Users\death\anaconda3\envs\openmm\lib\site-packages\pandas\plotting\_matplotlib\m
isc.py:101: UserWarning: Attempting to set identical bottom == top == 52.0 results i
n singular transformations; automatically expanding.
    ax.set_ylim(boundaries_list[i])
c:\Users\death\anaconda3\envs\openmm\lib\site-packages\pandas\plotting\_matplotlib\m
isc.py:91: UserWarning: Attempting to set identical left == right == 52.0 results in
singular transformations; automatically expanding.
    ax.set_xlim(boundaries_list[i])
```

```

Out[ ]: array([[<AxesSubplot:xlabel='n', ylabel='n'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='n'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='n'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='n'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='n'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='n'>,
<AxesSubplot:xlabel='crack_use', ylabel='n'>,
<AxesSubplot:xlabel='heroin_use', ylabel='n'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='n'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='n'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='n'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='n'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='n'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='n'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='n'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='n'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='n'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='n'>,
<AxesSubplot:xlabel='meth_use', ylabel='n'>,
<AxesSubplot:xlabel='sedative_use', ylabel='n'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='n'>],
[<AxesSubplot:xlabel='n', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='alcohol_use'>],
[<AxesSubplot:xlabel='n', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='crack_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='heroin_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='alcohol_frequenc
y'>,

```

```

<AxesSubplot:xlabel='oxycontin_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='meth_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='sedative_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='alcohol_frequency'>],
[<AxesSubplot:xlabel='n', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='marijuana_use'>],
[<AxesSubplot:xlabel='n', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='crack_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='heroin_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='marijuana_frequenc
y'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='marijuana_frequenc
y'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='marijuana_frequenc
y'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='meth_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='sedative_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='marijuana_frequency'>],
[<AxesSubplot:xlabel='n', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='cocaine_use'>,

```

```

<AxesSubplot:xlabel='marijuana_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='cocaine_use'>],
[<AxesSubplot:xlabel='n', ylabel='crack_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='crack_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='crack_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='crack_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='crack_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='crack_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='crack_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='crack_use'>],
[<AxesSubplot:xlabel='n', ylabel='heroin_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='heroin_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='heroin_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='heroin_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='heroin_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='heroin_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='heroin_use'>,

```

```

<AxesSubplot:xlabel='stimulant_frequency', ylabel='heroin_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='heroin_use'>],
[<AxesSubplot:xlabel='n', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='hallucinogen_use'>],
[<AxesSubplot:xlabel='n', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='hallucinogen_frequenc
y'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='crack_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='heroin_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='hallucinogen_frequen
cy'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='hallucinogen_frequen
cy'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='hallucinogen_frequen
cy'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='hallucinogen_frequenc
y'>,
<AxesSubplot:xlabel='meth_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='sedative_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='hallucinogen_frequenc
y'>],
[<AxesSubplot:xlabel='n', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='inhalant_use'>,

```

```

<AxesSubplot:xlabel='marijuana_frequency', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='inhalant_use'>],
[<AxesSubplot:xlabel='n', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='pain_releiver_us
e'>,

<AxesSubplot:xlabel='oxycontin_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='pain_releiver_use'>],
[<AxesSubplot:xlabel='n', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='pain_releiver_frequenc
y'>,

<AxesSubplot:xlabel='marijuana_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='pain_releiver_frequenc
y'>,

<AxesSubplot:xlabel='cocaine_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='crack_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='heroin_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='pain_releiver_frequen
cy'>,

<AxesSubplot:xlabel='inhalant_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='pain_releiver_frequenc
y'>,

<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='pain_releiver_frequ

```

```

ency'>,
  <AxesSubplot:xlabel='oxycontin_use', ylabel='pain_releiver_frequency'>,
  <AxesSubplot:xlabel='tranquilizer_use', ylabel='pain_releiver_frequency'>,
  <AxesSubplot:xlabel='tranquilizer_frequency', ylabel='pain_releiver_frequen
ncy'>,
  <AxesSubplot:xlabel='stimulant_use', ylabel='pain_releiver_frequency'>,
  <AxesSubplot:xlabel='stimulant_frequency', ylabel='pain_releiver_frequenc
y'>,
  <AxesSubplot:xlabel='meth_use', ylabel='pain_releiver_frequency'>,
  <AxesSubplot:xlabel='sedative_use', ylabel='pain_releiver_frequency'>,
  <AxesSubplot:xlabel='sedative_frequency', ylabel='pain_releiver_frequenc
y'>],
[<AxesSubplot:xlabel='n', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='alcohol_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='alcohol_frequency', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='marijuana_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='marijuana_frequency', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='cocaine_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='crack_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='heroin_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='hallucinogen_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='hallucinogen_frequency', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='inhalant_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='pain_releiver_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='pain_releiver_frequency', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='oxycontin_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='tranquilizer_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='tranquilizer_frequency', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='stimulant_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='stimulant_frequency', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='meth_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='sedative_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='sedative_frequency', ylabel='oxycontin_use'>],
[<AxesSubplot:xlabel='n', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='alcohol_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='alcohol_frequency', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='marijuana_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='marijuana_frequency', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='cocaine_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='crack_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='heroin_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='hallucinogen_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='hallucinogen_frequency', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='inhalant_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='pain_releiver_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='pain_releiver_frequency', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='oxycontin_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='tranquilizer_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='tranquilizer_frequency', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='stimulant_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='stimulant_frequency', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='meth_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='sedative_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='sedative_frequency', ylabel='tranquilizer_use'>],
[<AxesSubplot:xlabel='n', ylabel='tranquilizer_frequency'>,
  <AxesSubplot:xlabel='alcohol_use', ylabel='tranquilizer_frequency'>,

```

```

<AxesSubplot:xlabel='alcohol_frequency', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='tranquilizer_frequenc
y'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='crack_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='heroin_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='tranquilizer_frequen
cy'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='tranquilizer_frequen
cy'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='tranquilizer_frequen
cy'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='tranquilizer_frequenc
y'>,
<AxesSubplot:xlabel='meth_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='sedative_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='tranquilizer_frequenc
y'>],
[<AxesSubplot:xlabel='n', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='stimulant_use'>],
[<AxesSubplot:xlabel='n', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='crack_use', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='heroin_use', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='stimulant_frequenc

```

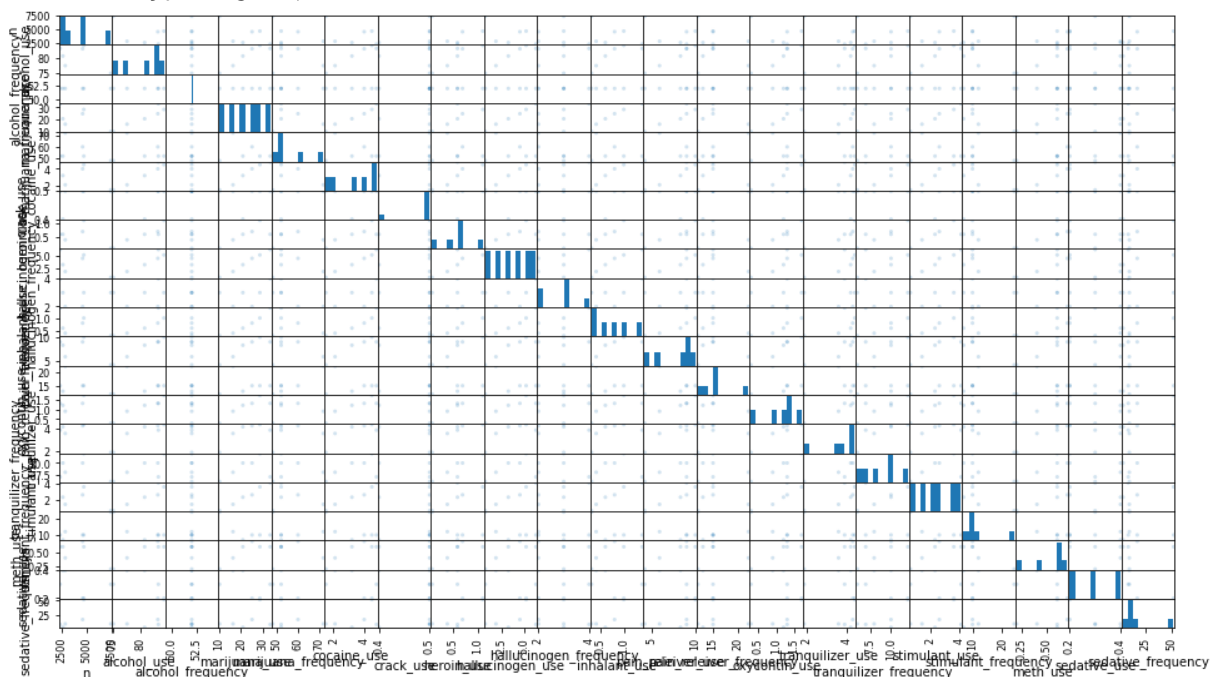


```

y'>,
    <AxesSubplot:xlabel='inhalant_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='pain_releiver_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='pain_releiver_frequency', ylabel='stimulant_frequenc
y'>,
    <AxesSubplot:xlabel='oxycontin_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='tranquilizer_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='tranquilizer_frequency', ylabel='stimulant_frequenc
y'>,
    <AxesSubplot:xlabel='stimulant_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='stimulant_frequency', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='meth_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='sedative_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='sedative_frequency', ylabel='stimulant_frequency'>],
[<AxesSubplot:xlabel='n', ylabel='meth_use'>,
    <AxesSubplot:xlabel='alcohol_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='alcohol_frequency', ylabel='meth_use'>,
    <AxesSubplot:xlabel='marijuana_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='marijuana_frequency', ylabel='meth_use'>,
    <AxesSubplot:xlabel='cocaine_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='crack_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='heroin_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='hallucinogen_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='hallucinogen_frequency', ylabel='meth_use'>,
    <AxesSubplot:xlabel='inhalant_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='pain_releiver_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='pain_releiver_frequency', ylabel='meth_use'>,
    <AxesSubplot:xlabel='oxycontin_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='tranquilizer_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='tranquilizer_frequency', ylabel='meth_use'>,
    <AxesSubplot:xlabel='stimulant_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='stimulant_frequency', ylabel='meth_use'>,
    <AxesSubplot:xlabel='meth_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='sedative_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='sedative_frequency', ylabel='meth_use'>],
[<AxesSubplot:xlabel='n', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='alcohol_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='alcohol_frequency', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='marijuana_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='marijuana_frequency', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='cocaine_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='crack_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='heroin_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='hallucinogen_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='hallucinogen_frequency', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='inhalant_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='pain_releiver_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='pain_releiver_frequency', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='oxycontin_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='tranquilizer_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='tranquilizer_frequency', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='stimulant_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='stimulant_frequency', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='meth_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='sedative_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='sedative_frequency', ylabel='sedative_use'>],

```

```
[<AxesSubplot:xlabel='n', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='crack_use', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='heroin_use', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='sedative_freque
n'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='sedative_freque
n'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='sedative_freque
n'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='meth_use', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='sedative_use', ylabel='sedative_frequency'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='sedative_frequency'>]],
dtype=object)
```



Example: Extracting Data from Kaggle.com

Let's now show you how to grab some data from Kaggle.com. Heads up, **you'll need a Kaggle profile for this.**

Kaggle has a whole section dedicated to public datasets, [here](#). Today, we're downloading the famous iris dataset from [this link](#).

Instructions

1. Go to this [link](#).
2. Download and unzip the zip file. Then upload the `Iris.csv` to Google Colab page.

```
In [ ]: #Read the Iris.csv file using pandas
iris_data = pd.read_csv("Iris.csv")
```

```
In [ ]: #Examine the first five and the last five rows.
print(iris_data.head(5))
print(iris_data.tail(5))
```

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	\
145	146	6.7	3.0	5.2	2.3	
146	147	6.3	2.5	5.0	1.9	
147	148	6.5	3.0	5.2	2.0	
148	149	6.2	3.4	5.4	2.3	
149	150	5.9	3.0	5.1	1.8	

	Species
145	Iris-virginica
146	Iris-virginica
147	Iris-virginica
148	Iris-virginica
149	Iris-virginica

```
In [ ]: #Examine a random sample
print(iris_data.sample())
```

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	\
139	140	6.9	3.1	5.4	2.1	

	Species
139	Iris-virginica

There is no inconsistency because there are no null values!

```
In [ ]: #Explore the data and check if there is any inconsistency in the Dataset.
print(iris_data.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 6 columns):
 #   Column          Non-Null Count  Dtype
---  -
 0   Id              150 non-null   int64
 1   SepalLengthCm   150 non-null   float64
 2   SepalWidthCm    150 non-null   float64
 3   PetalLengthCm   150 non-null   float64
 4   PetalWidthCm    150 non-null   float64
 5   Species         150 non-null   object
dtypes: float64(4), int64(1), object(1)
memory usage: 7.2+ KB
None
```

Exploratory Data Analysis

Exploratory Data Analysis (EDA) is a critical step in data analysis that identifies general patterns and relations in the data.

Lets start with the For example, we use scatter plots that show pairwise relationships between several variables.

Now lets try to plot the data and have some ideas about the relation between different columns.

```
In [ ]: from pandas.plotting import scatter_matrix
import matplotlib.pyplot as plt

scatter_matrix(data, figsize=(14,14))

plt.show()
```

```

Out[ ]: array([[<AxesSubplot:xlabel='n', ylabel='n'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='n'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='n'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='n'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='n'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='n'>,
<AxesSubplot:xlabel='crack_use', ylabel='n'>,
<AxesSubplot:xlabel='heroin_use', ylabel='n'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='n'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='n'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='n'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='n'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='n'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='n'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='n'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='n'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='n'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='n'>,
<AxesSubplot:xlabel='meth_use', ylabel='n'>,
<AxesSubplot:xlabel='sedative_use', ylabel='n'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='n'>],
[<AxesSubplot:xlabel='n', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='alcohol_use'>],
[<AxesSubplot:xlabel='n', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='crack_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='heroin_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='alcohol_frequency'>,
y'>,

```

```

<AxesSubplot:xlabel='oxycontin_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='meth_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='sedative_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='alcohol_frequency'>],
[<AxesSubplot:xlabel='n', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='marijuana_use'>],
[<AxesSubplot:xlabel='n', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='crack_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='heroin_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='marijuana_frequenc
y'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='marijuana_frequenc
y'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='marijuana_frequenc
y'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='meth_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='sedative_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='marijuana_frequency'>],
[<AxesSubplot:xlabel='n', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='cocaine_use'>,

```

```

<AxesSubplot:xlabel='marijuana_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='cocaine_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='cocaine_use'>],
[<AxesSubplot:xlabel='n', ylabel='crack_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='crack_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='crack_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='crack_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='crack_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='crack_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='crack_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='crack_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='crack_use'>],
[<AxesSubplot:xlabel='n', ylabel='heroin_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='heroin_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='heroin_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='heroin_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='heroin_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='heroin_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='heroin_use'>,

```

```

<AxesSubplot:xlabel='stimulant_frequency', ylabel='heroin_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='heroin_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='heroin_use'>],
[<AxesSubplot:xlabel='n', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='hallucinogen_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='hallucinogen_use'>],
[<AxesSubplot:xlabel='n', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='hallucinogen_frequenc
y'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='crack_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='heroin_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='hallucinogen_frequen
cy'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='hallucinogen_frequen
cy'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='hallucinogen_frequen
cy'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='hallucinogen_frequenc
y'>,
<AxesSubplot:xlabel='meth_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='sedative_use', ylabel='hallucinogen_frequency'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='hallucinogen_frequenc
y'>],
[<AxesSubplot:xlabel='n', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='inhalant_use'>,

```



```

<AxesSubplot:xlabel='marijuana_frequency', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='inhalant_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='inhalant_use'>],
[<AxesSubplot:xlabel='n', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='pain_releiver_us
e'>,

<AxesSubplot:xlabel='oxycontin_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='pain_releiver_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='pain_releiver_use'>],
[<AxesSubplot:xlabel='n', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='pain_releiver_frequenc
y'>,

<AxesSubplot:xlabel='marijuana_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='pain_releiver_frequenc
y'>,

<AxesSubplot:xlabel='cocaine_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='crack_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='heroin_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='pain_releiver_frequen
cy'>,

<AxesSubplot:xlabel='inhalant_use', ylabel='pain_releiver_frequency'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='pain_releiver_frequenc
y'>,

<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='pain_releiver_frequ

```

```

ency'>,
  <AxesSubplot:xlabel='oxycontin_use', ylabel='pain_releiver_frequency'>,
  <AxesSubplot:xlabel='tranquilizer_use', ylabel='pain_releiver_frequency'>,
  <AxesSubplot:xlabel='tranquilizer_frequency', ylabel='pain_releiver_frequen
ncy'>,
  <AxesSubplot:xlabel='stimulant_use', ylabel='pain_releiver_frequency'>,
  <AxesSubplot:xlabel='stimulant_frequency', ylabel='pain_releiver_frequenc
y'>,
  <AxesSubplot:xlabel='meth_use', ylabel='pain_releiver_frequency'>,
  <AxesSubplot:xlabel='sedative_use', ylabel='pain_releiver_frequency'>,
  <AxesSubplot:xlabel='sedative_frequency', ylabel='pain_releiver_frequenc
y'>],
[<AxesSubplot:xlabel='n', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='alcohol_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='alcohol_frequency', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='marijuana_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='marijuana_frequency', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='cocaine_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='crack_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='heroin_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='hallucinogen_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='hallucinogen_frequency', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='inhalant_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='pain_releiver_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='pain_releiver_frequency', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='oxycontin_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='tranquilizer_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='tranquilizer_frequency', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='stimulant_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='stimulant_frequency', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='meth_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='sedative_use', ylabel='oxycontin_use'>,
  <AxesSubplot:xlabel='sedative_frequency', ylabel='oxycontin_use'>],
[<AxesSubplot:xlabel='n', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='alcohol_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='alcohol_frequency', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='marijuana_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='marijuana_frequency', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='cocaine_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='crack_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='heroin_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='hallucinogen_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='hallucinogen_frequency', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='inhalant_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='pain_releiver_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='pain_releiver_frequency', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='oxycontin_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='tranquilizer_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='tranquilizer_frequency', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='stimulant_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='stimulant_frequency', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='meth_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='sedative_use', ylabel='tranquilizer_use'>,
  <AxesSubplot:xlabel='sedative_frequency', ylabel='tranquilizer_use'>],
[<AxesSubplot:xlabel='n', ylabel='tranquilizer_frequency'>,
  <AxesSubplot:xlabel='alcohol_use', ylabel='tranquilizer_frequency'>,

```

```

<AxesSubplot:xlabel='alcohol_frequency', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='tranquilizer_frequenc
y'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='crack_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='heroin_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='tranquilizer_frequen
cy'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='tranquilizer_frequen
cy'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='tranquilizer_frequen
cy'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='tranquilizer_frequenc
y'>,
<AxesSubplot:xlabel='meth_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='sedative_use', ylabel='tranquilizer_frequency'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='tranquilizer_frequenc
y'>],
[<AxesSubplot:xlabel='n', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='crack_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='heroin_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='inhalant_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='pain_releiver_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='pain_releiver_frequency', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='oxycontin_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='tranquilizer_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='tranquilizer_frequency', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='stimulant_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='stimulant_frequency', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='meth_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='sedative_use', ylabel='stimulant_use'>,
<AxesSubplot:xlabel='sedative_frequency', ylabel='stimulant_use'>],
[<AxesSubplot:xlabel='n', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='alcohol_use', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='cocaine_use', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='crack_use', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='heroin_use', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='hallucinogen_use', ylabel='stimulant_frequency'>,
<AxesSubplot:xlabel='hallucinogen_frequency', ylabel='stimulant_frequenc

```

```

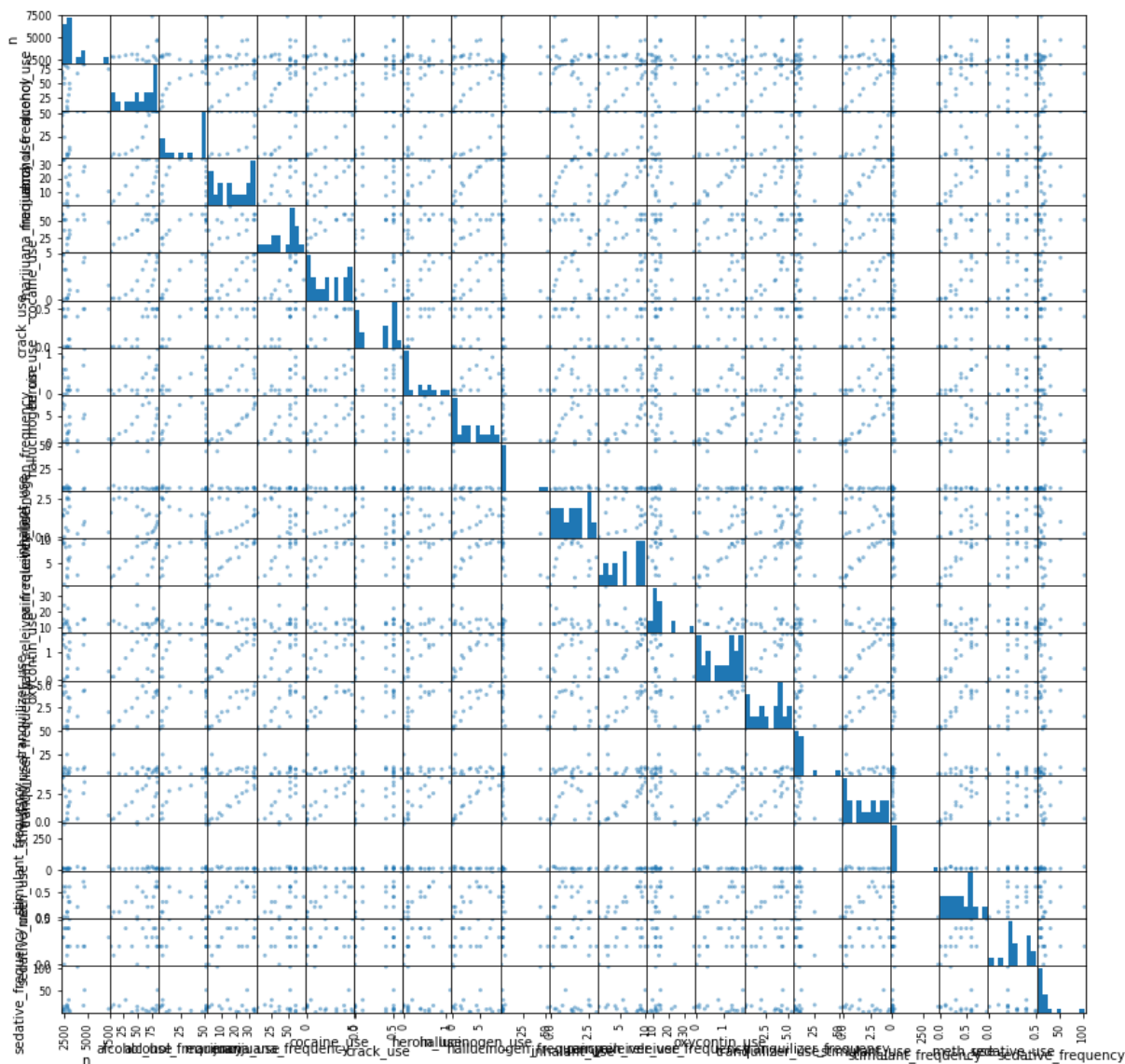
y'>,
    <AxesSubplot:xlabel='inhalant_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='pain_releiver_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='pain_releiver_frequency', ylabel='stimulant_frequenc
y'>,
    <AxesSubplot:xlabel='oxycontin_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='tranquilizer_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='tranquilizer_frequency', ylabel='stimulant_frequenc
y'>,
    <AxesSubplot:xlabel='stimulant_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='stimulant_frequency', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='meth_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='sedative_use', ylabel='stimulant_frequency'>,
    <AxesSubplot:xlabel='sedative_frequency', ylabel='stimulant_frequency'>],
[<AxesSubplot:xlabel='n', ylabel='meth_use'>,
    <AxesSubplot:xlabel='alcohol_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='alcohol_frequency', ylabel='meth_use'>,
    <AxesSubplot:xlabel='marijuana_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='marijuana_frequency', ylabel='meth_use'>,
    <AxesSubplot:xlabel='cocaine_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='crack_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='heroin_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='hallucinogen_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='hallucinogen_frequency', ylabel='meth_use'>,
    <AxesSubplot:xlabel='inhalant_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='pain_releiver_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='pain_releiver_frequency', ylabel='meth_use'>,
    <AxesSubplot:xlabel='oxycontin_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='tranquilizer_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='tranquilizer_frequency', ylabel='meth_use'>,
    <AxesSubplot:xlabel='stimulant_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='stimulant_frequency', ylabel='meth_use'>,
    <AxesSubplot:xlabel='meth_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='sedative_use', ylabel='meth_use'>,
    <AxesSubplot:xlabel='sedative_frequency', ylabel='meth_use'>],
[<AxesSubplot:xlabel='n', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='alcohol_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='alcohol_frequency', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='marijuana_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='marijuana_frequency', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='cocaine_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='crack_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='heroin_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='hallucinogen_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='hallucinogen_frequency', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='inhalant_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='pain_releiver_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='pain_releiver_frequency', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='oxycontin_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='tranquilizer_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='tranquilizer_frequency', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='stimulant_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='stimulant_frequency', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='meth_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='sedative_use', ylabel='sedative_use'>,
    <AxesSubplot:xlabel='sedative_frequency', ylabel='sedative_use'>],

```

```

[<AxesSubplot:xlabel='n', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='alcohol_use', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='alcohol_frequency', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='marijuana_use', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='marijuana_frequency', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='cocaine_use', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='crack_use', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='heroin_use', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='hallucinogen_use', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='hallucinogen_frequency', ylabel='sedative_freque
y'>,
 <AxesSubplot:xlabel='inhalant_use', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='pain_releiver_use', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='pain_releiver_frequency', ylabel='sedative_freque
y'>,
 <AxesSubplot:xlabel='oxycontin_use', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='tranquilizer_use', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='tranquilizer_frequency', ylabel='sedative_freque
y'>,
 <AxesSubplot:xlabel='stimulant_use', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='stimulant_frequency', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='meth_use', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='sedative_use', ylabel='sedative_frequency'>,
 <AxesSubplot:xlabel='sedative_frequency', ylabel='sedative_frequency'>]],
dtype=object)

```



When we're training algorithms, we need to pay attention to how many features we've got and how they relate to each other. If we've got a bunch of features that are all tangled up (correlated), using them all can mess with our accuracy. So, we've got to be smart about picking the right ones. But how?

In machine learning and data analysis, understanding feature correlations is important because highly correlated features can sometimes provide redundant or overlapping information to algorithms, which can lead to suboptimal model performance. Therefore, feature selection or feature engineering techniques are used to choose the most relevant and uncorrelated features for modeling.

Correlation We say two or more features (variables) in a dataset are correlated if there is a statistical relationship or association between them. When features are correlated, changes in one feature may be related to changes in another.

For example, if you have two features in a dataset, such as "temperature" and "ice cream sales," and you notice that as the temperature increases, ice cream sales also tend to

increase, these two features can be considered correlated. In simpler terms, when one goes up, the other often goes up as well, or vice versa. In most data, pairwise correlations may not provide enough insights, and multivariate approaches to exploratory data analyses are recommended.

```
In [ ]: # use .corr() to find pairwise correlation in CSV file of example 1  
        print(data.corr())
```

	n	alcohol_use	alcohol_frequency	\
n	1.000000	0.310512	0.313431	
alcohol_use	0.310512	1.000000	0.897889	
alcohol_frequency	0.313431	0.897889	1.000000	
marijuana_use	-0.169007	0.594165	0.260219	
marijuana_frequency	0.135597	0.920425	0.818769	
cocaine_use	-0.021565	0.773458	0.577601	
crack_use	0.281797	0.877378	0.785050	
heroin_use	0.007798	0.677614	0.546194	
hallucinogen_use	-0.295302	0.463702	0.182184	
hallucinogen_frequency	0.008743	-0.346643	-0.174622	
inhalant_use	-0.368395	-0.648248	-0.864153	
pain_releiver_use	-0.083587	0.617523	0.280704	
pain_releiver_frequency	-0.112076	-0.263721	0.007087	
oxycontin_use	-0.166039	0.589219	0.251554	
tranquilizer_use	-0.071086	0.735785	0.448543	
tranquilizer_frequency	-0.102301	-0.550501	-0.415026	
stimulant_use	-0.201621	0.582241	0.291766	
stimulant_frequency	-0.114792	-0.028115	0.260304	
meth_use	-0.063571	0.682531	0.425153	
sedative_use	-0.119997	0.318268	0.029867	
sedative_frequency	0.256701	0.091251	0.201423	

	marijuana_use	marijuana_frequency	cocaine_use	\
n	-0.169007	0.135597	-0.021565	
alcohol_use	0.594165	0.920425	0.773458	
alcohol_frequency	0.260219	0.818769	0.577601	
marijuana_use	1.000000	0.615610	0.883789	
marijuana_frequency	0.615610	1.000000	0.742598	
cocaine_use	0.883789	0.742598	1.000000	
crack_use	0.620670	0.892077	0.835727	
heroin_use	0.711613	0.631789	0.899944	
hallucinogen_use	0.959353	0.528733	0.855668	
hallucinogen_frequency	-0.478286	-0.380430	-0.368402	
inhalant_use	0.127336	-0.577792	-0.214609	
pain_releiver_use	0.966226	0.612254	0.906638	
pain_releiver_frequency	-0.506892	-0.285292	-0.273908	
oxycontin_use	0.971733	0.609618	0.865626	
tranquilizer_use	0.935199	0.768339	0.920430	
tranquilizer_frequency	-0.401769	-0.591449	-0.294905	
stimulant_use	0.968096	0.561246	0.923954	
stimulant_frequency	-0.389145	-0.069454	-0.309286	
meth_use	0.870446	0.664280	0.889353	
sedative_use	0.616766	0.399124	0.461914	
sedative_frequency	-0.277144	0.048777	-0.159997	

	crack_use	heroin_use	hallucinogen_use	\
n	0.281797	0.007798	-0.295302	
alcohol_use	0.877378	0.677614	0.463702	
alcohol_frequency	0.785050	0.546194	0.182184	
marijuana_use	0.620670	0.711613	0.959353	
marijuana_frequency	0.892077	0.631789	0.528733	
cocaine_use	0.835727	0.899944	0.855668	
crack_use	1.000000	0.750790	0.556222	
heroin_use	0.750790	1.000000	0.683674	
hallucinogen_use	0.556222	0.683674	1.000000	

hallucinogen_frequency	-0.198968	-0.302988	-0.427863
inhalant_use	-0.543499	-0.215961	0.173030
pain_releiver_use	0.652764	0.805868	0.909333
pain_releiver_frequency	-0.203711	-0.135961	-0.419961
oxycontin_use	0.603003	0.775321	0.916777
tranquilizer_use	0.772651	0.818583	0.862900
tranquilizer_frequency	-0.345668	-0.175464	-0.314281
stimulant_use	0.618008	0.808024	0.940008
stimulant_frequency	-0.294260	-0.277996	-0.319860
meth_use	0.694324	0.816620	0.780071
sedative_use	0.438421	0.333649	0.518700
sedative_frequency	0.091676	-0.023708	-0.312746

	hallucinogen_frequency	...	pain_releiver_use	\
n	0.008743	...	-0.083587	
alcohol_use	-0.346643	...	0.617523	
alcohol_frequency	-0.174622	...	0.280704	
marijuana_use	-0.478286	...	0.966226	
marijuana_frequency	-0.380430	...	0.612254	
cocaine_use	-0.368402	...	0.906638	
crack_use	-0.198968	...	0.652764	
heroin_use	-0.302988	...	0.805868	
hallucinogen_use	-0.427863	...	0.909333	
hallucinogen_frequency	1.000000	...	-0.488187	
inhalant_use	-0.137576	...	0.092045	
pain_releiver_use	-0.488187	...	1.000000	
pain_releiver_frequency	0.547363	...	-0.468617	
oxycontin_use	-0.442181	...	0.976445	
tranquilizer_use	-0.458016	...	0.959283	
tranquilizer_frequency	0.705734	...	-0.351856	
stimulant_use	-0.452345	...	0.961484	
stimulant_frequency	-0.103980	...	-0.471433	
meth_use	-0.432520	...	0.918051	
sedative_use	-0.231773	...	0.636666	
sedative_frequency	0.533095	...	-0.271246	

	pain_releiver_frequency	oxycontin_use	\
n	-0.112076	-0.166039	
alcohol_use	-0.263721	0.589219	
alcohol_frequency	0.007087	0.251554	
marijuana_use	-0.506892	0.971733	
marijuana_frequency	-0.285292	0.609618	
cocaine_use	-0.273908	0.865626	
crack_use	-0.203711	0.603003	
heroin_use	-0.135961	0.775321	
hallucinogen_use	-0.419961	0.916777	
hallucinogen_frequency	0.547363	-0.442181	
inhalant_use	-0.342676	0.110492	
pain_releiver_use	-0.468617	0.976445	
pain_releiver_frequency	1.000000	-0.474503	
oxycontin_use	-0.474503	1.000000	
tranquilizer_use	-0.414122	0.955081	
tranquilizer_frequency	0.699282	-0.355122	
stimulant_use	-0.409813	0.947492	
stimulant_frequency	0.324717	-0.406369	
meth_use	-0.438606	0.895790	

sedative_use	-0.378001	0.625929
sedative_frequency	-0.045213	-0.198329

	tranquilizer_use	tranquilizer_frequency	\
n	-0.071086	-0.102301	
alcohol_use	0.735785	-0.550501	
alcohol_frequency	0.448543	-0.415026	
marijuana_use	0.935199	-0.401769	
marijuana_frequency	0.768339	-0.591449	
cocaine_use	0.920430	-0.294905	
crack_use	0.772651	-0.345668	
heroin_use	0.818583	-0.175464	
hallucinogen_use	0.862900	-0.314281	
hallucinogen_frequency	-0.458016	0.705734	
inhalant_use	-0.103745	0.158140	
pain_releiver_use	0.959283	-0.351856	
pain_releiver_frequency	-0.414122	0.699282	
oxycontin_use	0.955081	-0.355122	
tranquilizer_use	1.000000	-0.388954	
tranquilizer_frequency	-0.388954	1.000000	
stimulant_use	0.909784	-0.327899	
stimulant_frequency	-0.384431	-0.180415	
meth_use	0.936270	-0.347500	
sedative_use	0.674618	-0.245645	
sedative_frequency	-0.224045	-0.014524	

	stimulant_use	stimulant_frequency	meth_use	\
n	-0.201621	-0.114792	-0.063571	
alcohol_use	0.582241	-0.028115	0.682531	
alcohol_frequency	0.291766	0.260304	0.425153	
marijuana_use	0.968096	-0.389145	0.870446	
marijuana_frequency	0.561246	-0.069454	0.664280	
cocaine_use	0.923954	-0.309286	0.889353	
crack_use	0.618008	-0.294260	0.694324	
heroin_use	0.808024	-0.277996	0.816620	
hallucinogen_use	0.940008	-0.319860	0.780071	
hallucinogen_frequency	-0.452345	-0.103980	-0.432520	
inhalant_use	0.108381	-0.419717	-0.059903	
pain_releiver_use	0.961484	-0.471433	0.918051	
pain_releiver_frequency	-0.409813	0.324717	-0.438606	
oxycontin_use	0.947492	-0.406369	0.895790	
tranquilizer_use	0.909784	-0.384431	0.936270	
tranquilizer_frequency	-0.327899	-0.180415	-0.347500	
stimulant_use	1.000000	-0.363358	0.894757	
stimulant_frequency	-0.363358	1.000000	-0.376278	
meth_use	0.894757	-0.376278	1.000000	
sedative_use	0.567956	-0.526289	0.663079	
sedative_frequency	-0.260968	-0.011796	-0.187589	

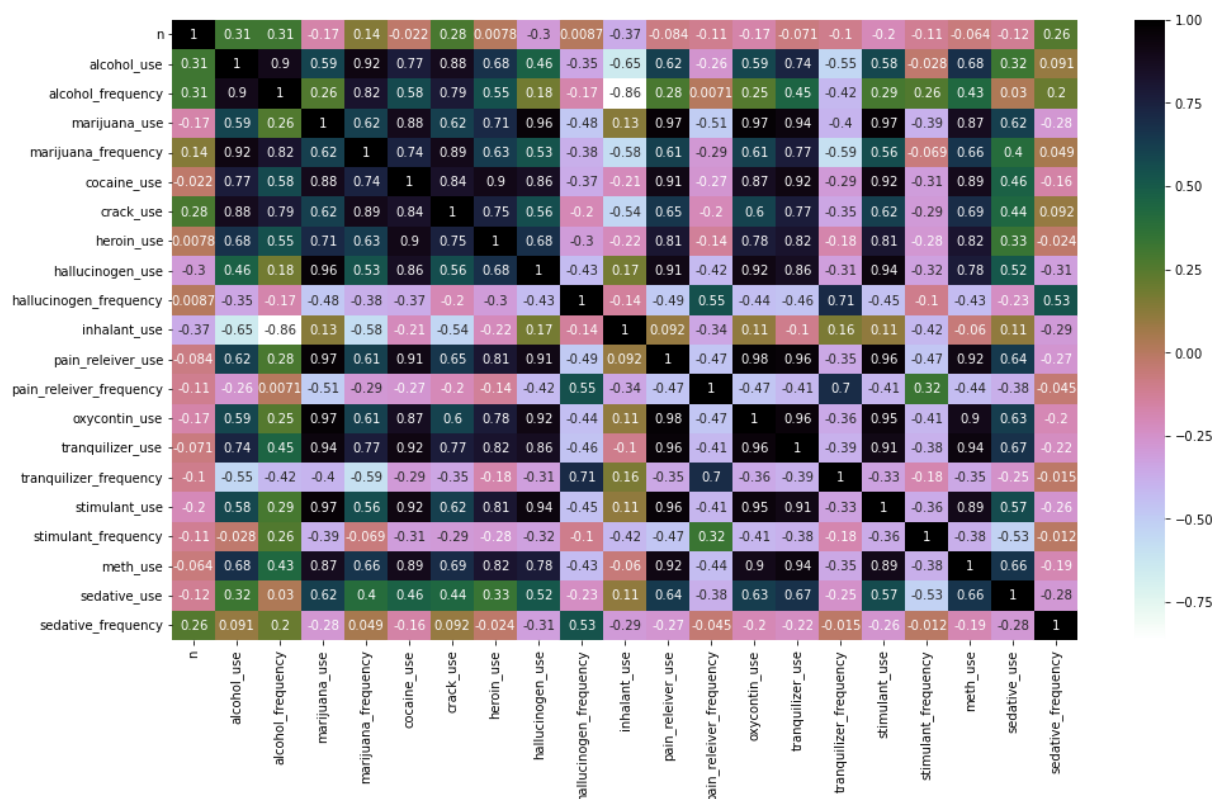
	sedative_use	sedative_frequency
n	-0.119997	0.256701
alcohol_use	0.318268	0.091251
alcohol_frequency	0.029867	0.201423
marijuana_use	0.616766	-0.277144
marijuana_frequency	0.399124	0.048777
cocaine_use	0.461914	-0.159997

crack_use	0.438421	0.091676
heroin_use	0.333649	-0.023708
hallucinogen_use	0.518700	-0.312746
hallucinogen_frequency	-0.231773	0.533095
inhalant_use	0.110611	-0.293189
pain_releiver_use	0.636666	-0.271246
pain_releiver_frequency	-0.378001	-0.045213
oxycontin_use	0.625929	-0.198329
tranquilizer_use	0.674618	-0.224045
tranquilizer_frequency	-0.245645	-0.014524
stimulant_use	0.567956	-0.260968
stimulant_frequency	-0.526289	-0.011796
meth_use	0.663079	-0.187589
sedative_use	1.000000	-0.284234
sedative_frequency	-0.284234	1.000000

[21 rows x 21 columns]

```
In [ ]: # run this code for a better view
plt.figure(figsize=(16,9))
import seaborn as sns

sns.heatmap(data.corr(), annot=True, cmap='cubehelix_r')
plt.show()
```



Let's say your project focuses on only alcohol, marijuana and cocaine. Create a table that specifically captures data on alcohol, marijuana, and cocaine.

```
In [ ]: # code here
data1 = data.iloc[0:15]
data1 = data1[['alcohol_use', 'alcohol_frequency', 'marijuana_use', 'marijuana_frequen
```

```
data1['cocaine_frequency'] = pd.to_numeric(data1['cocaine_frequency'])
print(data1)
```

	alcohol_use	alcohol_frequency	marijuana_use	marijuana_frequency	\
0	3.9	3.0	1.1	4.0	
1	8.5	6.0	3.4	15.0	
2	18.1	5.0	8.7	24.0	
3	29.2	6.0	14.5	25.0	
4	40.1	10.0	22.5	30.0	
5	49.3	13.0	28.0	36.0	
6	58.7	24.0	33.7	52.0	
7	64.6	36.0	33.4	60.0	
8	69.7	48.0	34.0	60.0	
9	83.2	52.0	33.0	52.0	
10	84.2	52.0	28.4	52.0	
11	83.1	52.0	24.9	60.0	
12	80.7	52.0	20.8	52.0	
13	77.5	52.0	16.4	72.0	
14	75.0	52.0	10.4	48.0	

	cocaine_frequency
0	5.0
1	1.0
2	5.5
3	4.0
4	7.0
5	5.0
6	5.0
7	5.5
8	8.0
9	5.0
10	5.0
11	6.0
12	5.0
13	8.0
14	15.0

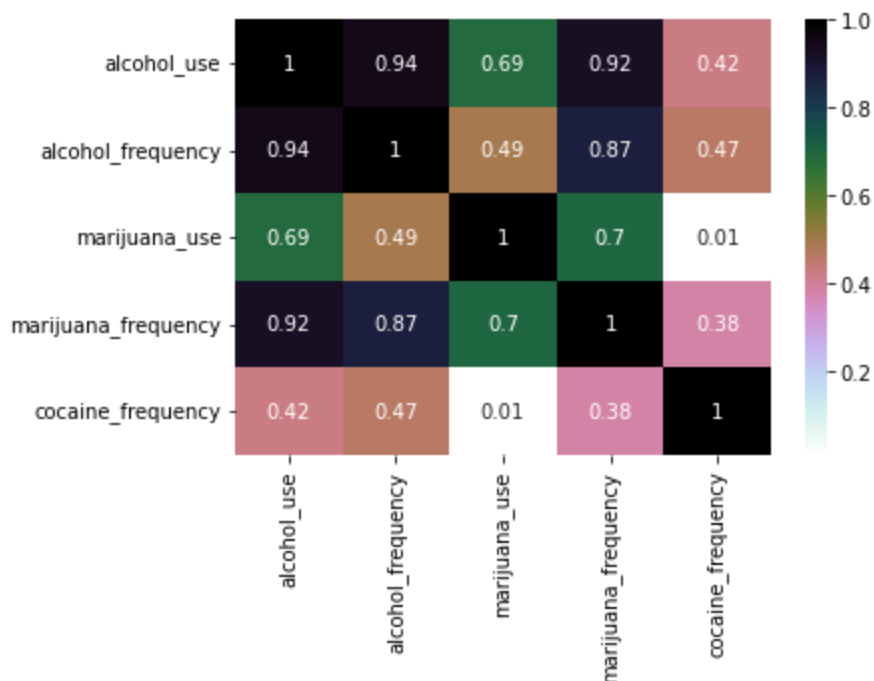
```
In [ ]: # 1- Find the pairwise correlation in this new table.
print(data1.corr())
```

	alcohol_use	alcohol_frequency	marijuana_use	\
alcohol_use	1.000000	0.944436	0.687316	
alcohol_frequency	0.944436	1.000000	0.494774	
marijuana_use	0.687316	0.494774	1.000000	
marijuana_frequency	0.919660	0.870130	0.701821	
cocaine_frequency	0.424428	0.465048	0.010256	

	marijuana_frequency	cocaine_frequency
alcohol_use	0.919660	0.424428
alcohol_frequency	0.870130	0.465048
marijuana_use	0.701821	0.010256
marijuana_frequency	1.000000	0.382617
cocaine_frequency	0.382617	1.000000

```
In [ ]: # 2- Plot the heatmap as above
sns.heatmap(data1.corr(), annot=True, cmap='cubehelix_r')
```

```
plt.show()
```



```
In [ ]: # 3- How do you interpret the result?
```

```
"""
```

Based on this heat graph, I can gather the following interpretations:

Marijuana frequency is heavily coorrelated with alcohol use and quite coorelated wi
nor alcohol. Alcohol use and frequency also seem heavily correlated with each other
"""

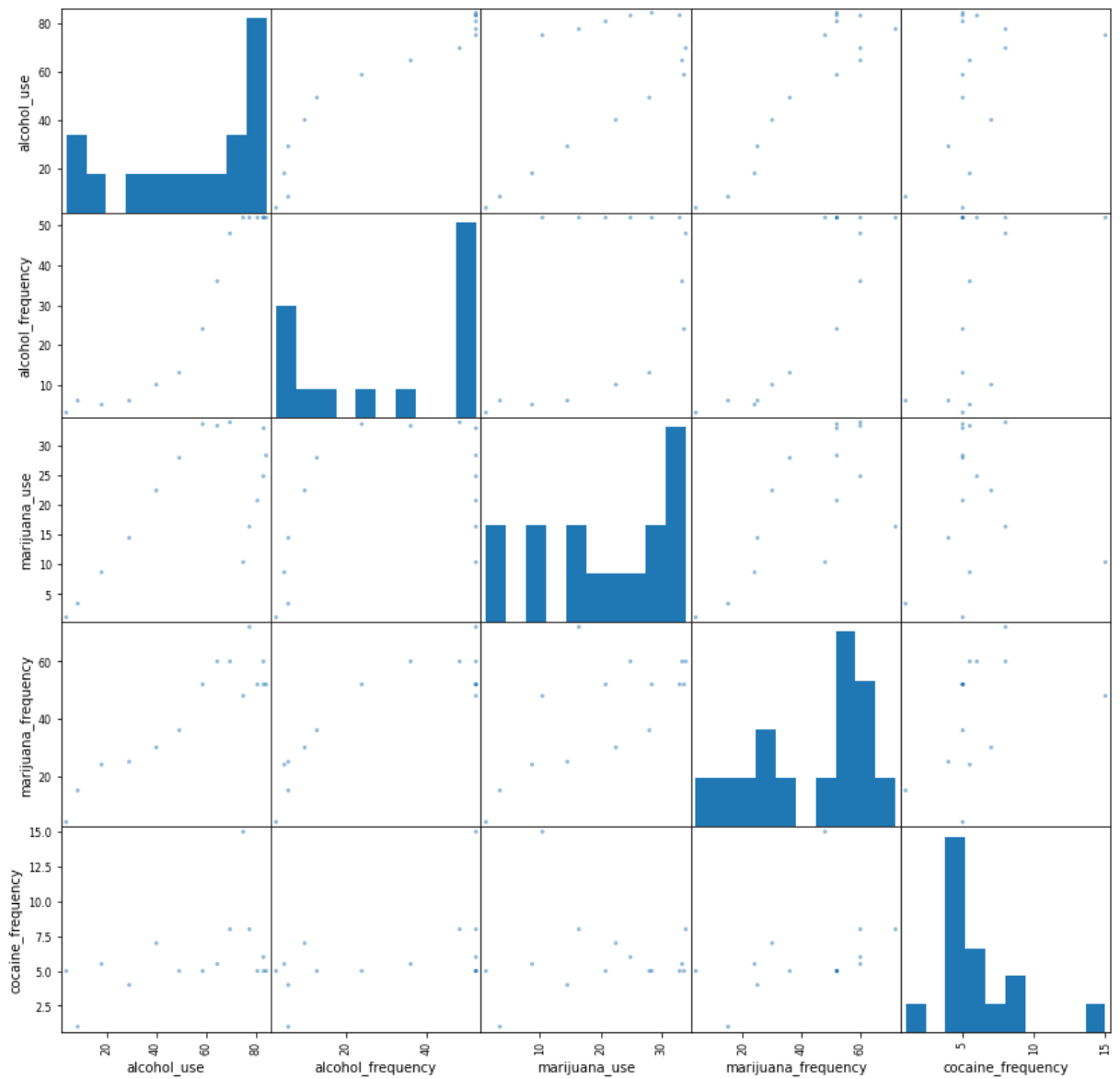
```
In [ ]: # 4- plot data using scatter matrix
```

```
scatter_matrix(data1, figsize=(14,14))
```

```

Out[ ]: array([[<AxesSubplot:xlabel='alcohol_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='alcohol_use'>,
<AxesSubplot:xlabel='cocaine_frequency', ylabel='alcohol_use'>],
[<AxesSubplot:xlabel='alcohol_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='alcohol_frequency'>,
<AxesSubplot:xlabel='cocaine_frequency', ylabel='alcohol_frequency'>],
[<AxesSubplot:xlabel='alcohol_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='marijuana_use'>,
<AxesSubplot:xlabel='cocaine_frequency', ylabel='marijuana_use'>],
[<AxesSubplot:xlabel='alcohol_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='marijuana_frequency'>,
<AxesSubplot:xlabel='cocaine_frequency', ylabel='marijuana_frequency'>],
[<AxesSubplot:xlabel='alcohol_use', ylabel='cocaine_frequency'>,
<AxesSubplot:xlabel='alcohol_frequency', ylabel='cocaine_frequency'>,
<AxesSubplot:xlabel='marijuana_use', ylabel='cocaine_frequency'>,
<AxesSubplot:xlabel='marijuana_frequency', ylabel='cocaine_frequency'>,
<AxesSubplot:xlabel='cocaine_frequency', ylabel='cocaine_frequency'>]],
dtype=object)

```

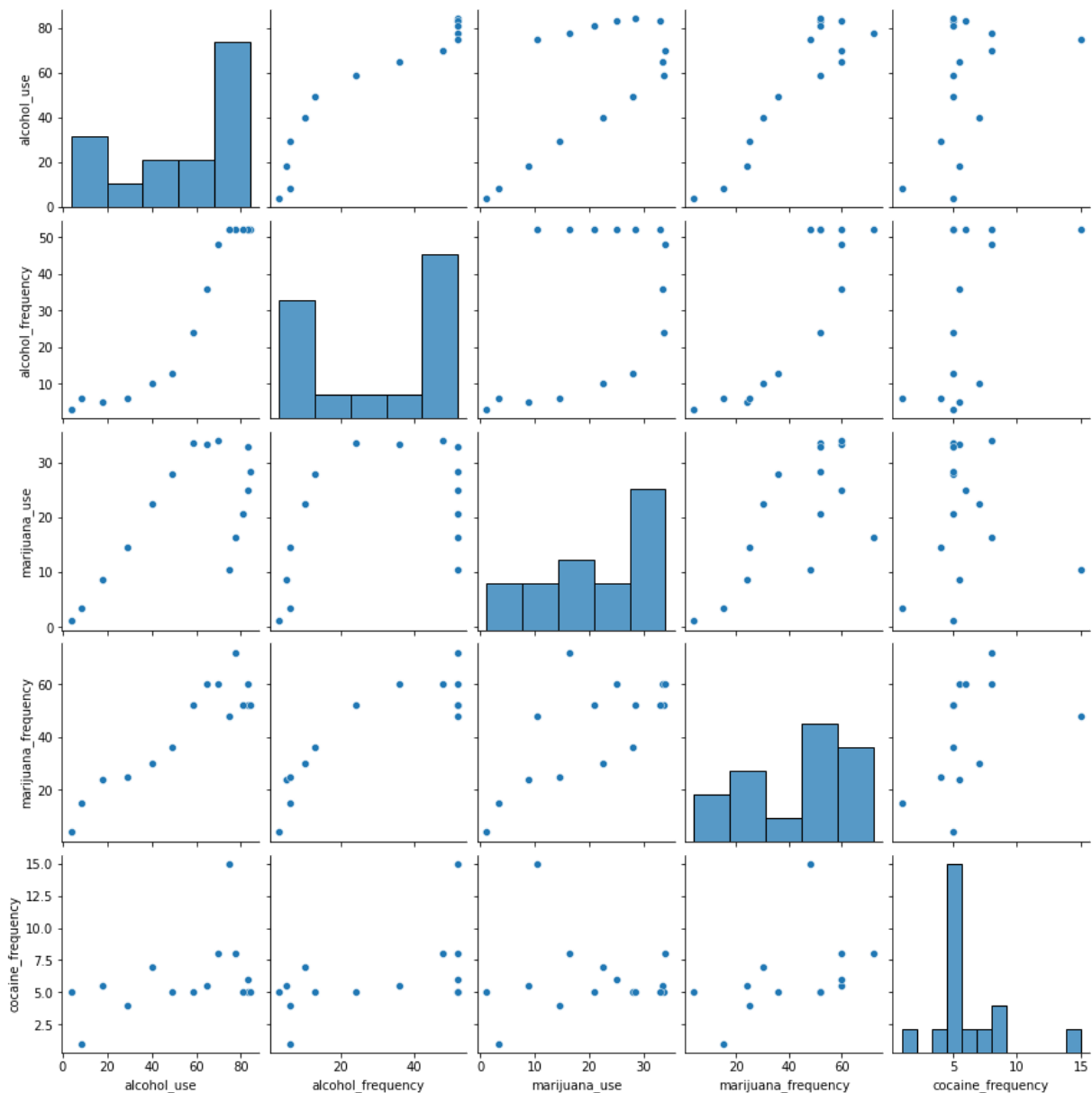


Let's try a different way of plotting: `sns.pairplot`. Check out its documentation to learn how to use it.

Documentation link: [Seaborn pairplot documentation](#)

```
In [ ]: #code her
sns.pairplot(data1)
```

```
Out[ ]: <seaborn.axisgrid.PairGrid at 0x1dc070646d0>
```



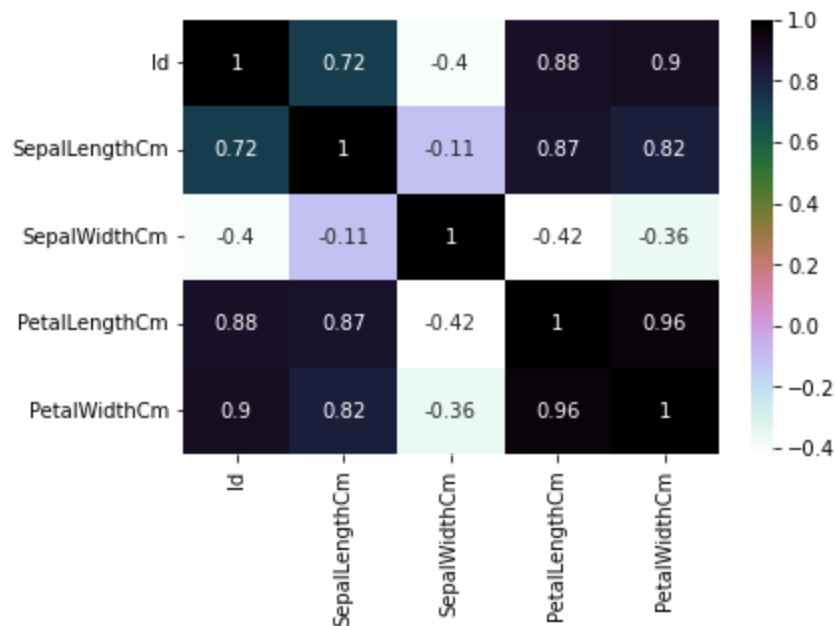
Follow the above steps for the CSV file in the second example.

```
In [ ]: # 1- Find the pairwise correlation in Iris.csv
print(iris_data.corr())
```

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	\
Id	1.000000	0.716676	-0.397729	0.882747	
SepalLengthCm	0.716676	1.000000	-0.109369	0.871754	
SepalWidthCm	-0.397729	-0.109369	1.000000	-0.420516	
PetalLengthCm	0.882747	0.871754	-0.420516	1.000000	
PetalWidthCm	0.899759	0.817954	-0.356544	0.962757	

	PetalWidthCm
Id	0.899759
SepalLengthCm	0.817954
SepalWidthCm	-0.356544
PetalLengthCm	0.962757
PetalWidthCm	1.000000


```
In [ ]: # 2- Plot the heatmap as above
sns.heatmap(iris_data.corr(), annot=True, cmap='cubehelix_r')
plt.show()
```



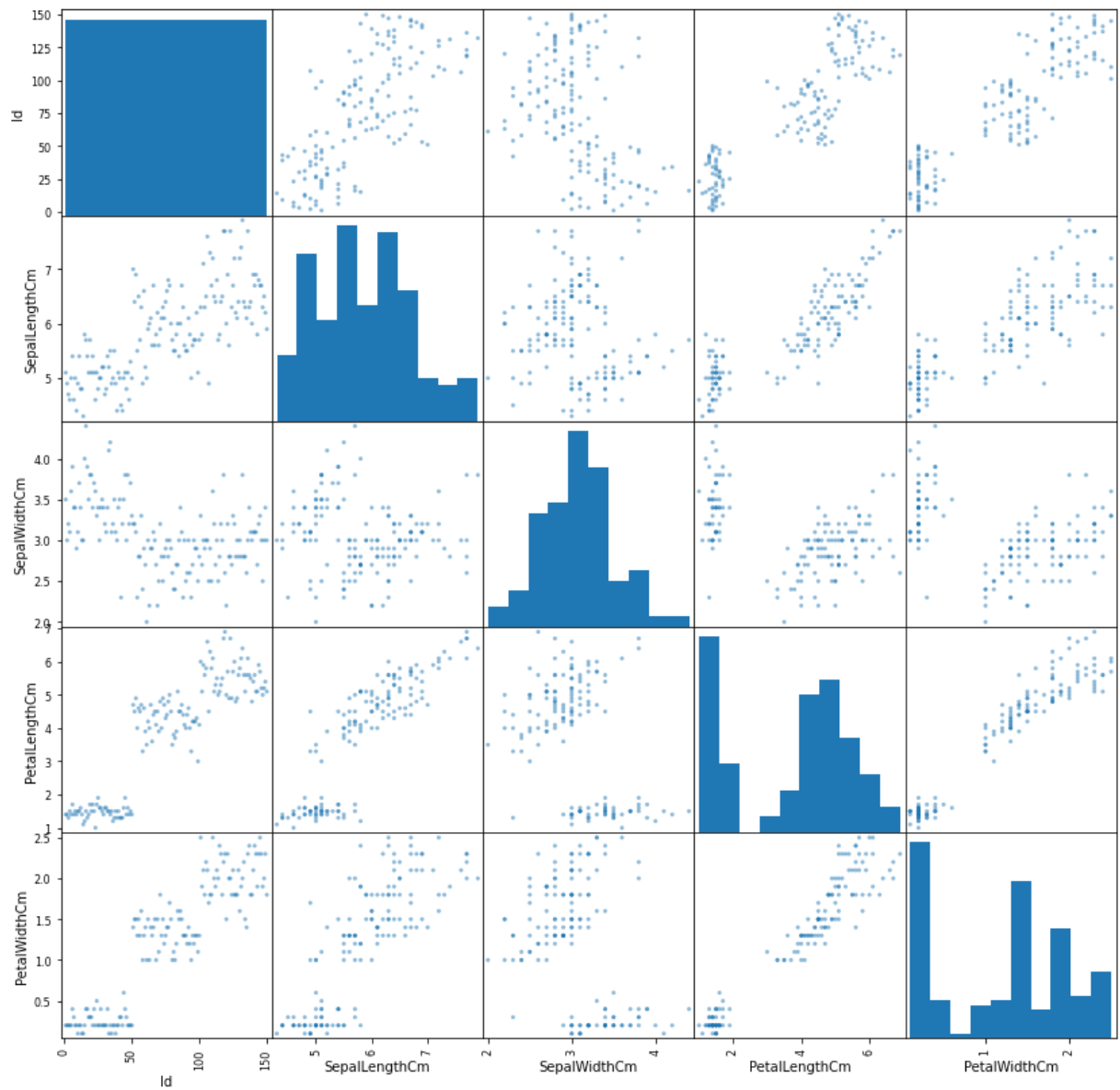
```
In [ ]: # 3- How do you interpret the result?
"""
There are a couple of correlations that we can see within this heat map:
Petal Width and Petal Length are heavily correlated with each; however, sepal width
positively correlated with Petal Length and Petal Width.
"""
```

```
In [ ]: # 4- Plot data using scatter matrix
scatter_matrix(iris_data, figsize=(14,14))
```

```

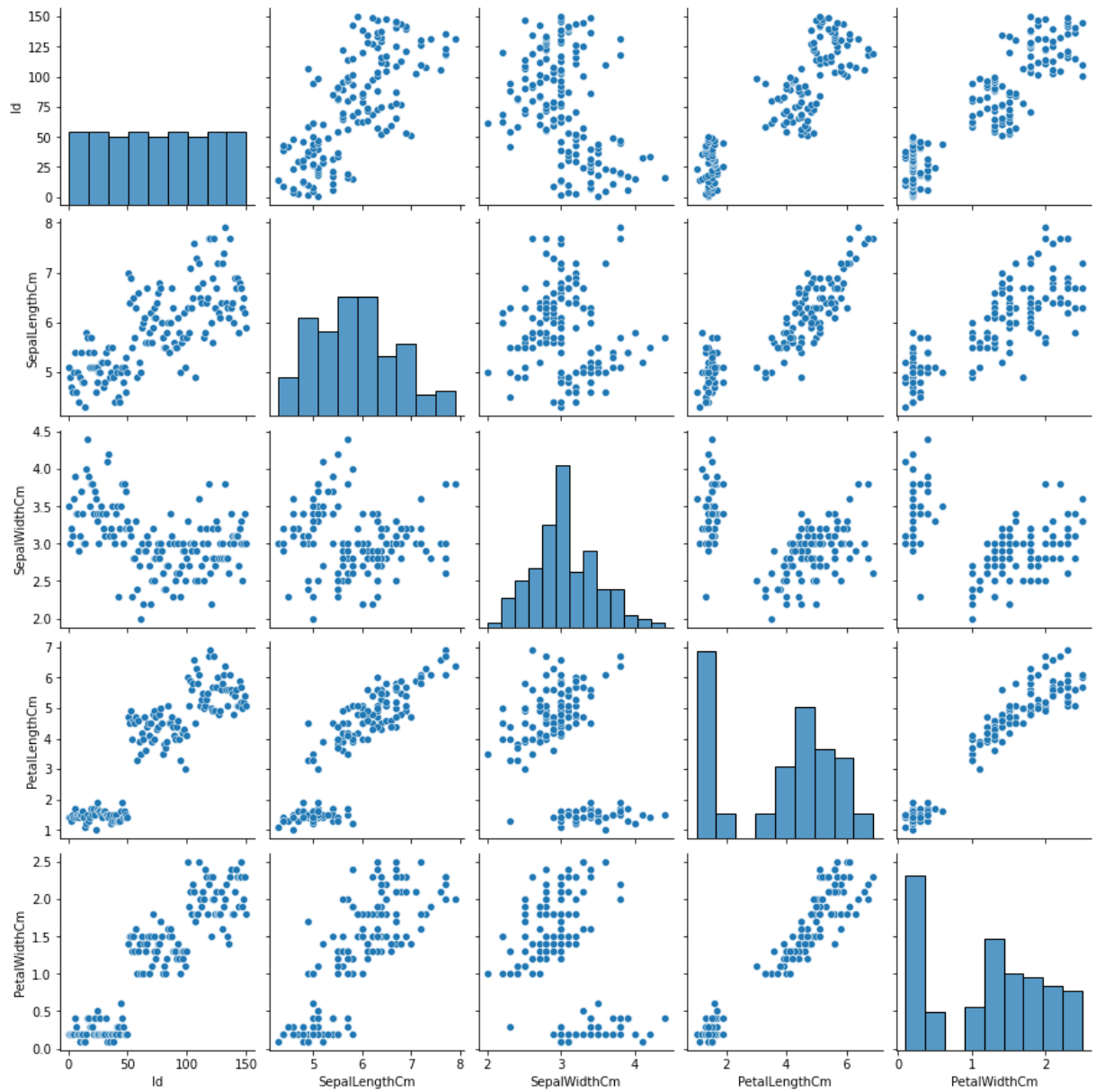
Out[ ]: array([[<AxesSubplot:xlabel='Id', ylabel='Id'>,
               <AxesSubplot:xlabel='SepalLengthCm', ylabel='Id'>,
               <AxesSubplot:xlabel='SepalWidthCm', ylabel='Id'>,
               <AxesSubplot:xlabel='PetalLengthCm', ylabel='Id'>,
               <AxesSubplot:xlabel='PetalWidthCm', ylabel='Id'>],
               [<AxesSubplot:xlabel='Id', ylabel='SepalLengthCm'>,
               <AxesSubplot:xlabel='SepalLengthCm', ylabel='SepalLengthCm'>,
               <AxesSubplot:xlabel='SepalWidthCm', ylabel='SepalLengthCm'>,
               <AxesSubplot:xlabel='PetalLengthCm', ylabel='SepalLengthCm'>,
               <AxesSubplot:xlabel='PetalWidthCm', ylabel='SepalLengthCm'>],
               [<AxesSubplot:xlabel='Id', ylabel='SepalWidthCm'>,
               <AxesSubplot:xlabel='SepalLengthCm', ylabel='SepalWidthCm'>,
               <AxesSubplot:xlabel='SepalWidthCm', ylabel='SepalWidthCm'>,
               <AxesSubplot:xlabel='PetalLengthCm', ylabel='SepalWidthCm'>,
               <AxesSubplot:xlabel='PetalWidthCm', ylabel='SepalWidthCm'>],
               [<AxesSubplot:xlabel='Id', ylabel='PetalLengthCm'>,
               <AxesSubplot:xlabel='SepalLengthCm', ylabel='PetalLengthCm'>,
               <AxesSubplot:xlabel='SepalWidthCm', ylabel='PetalLengthCm'>,
               <AxesSubplot:xlabel='PetalLengthCm', ylabel='PetalLengthCm'>,
               <AxesSubplot:xlabel='PetalWidthCm', ylabel='PetalLengthCm'>],
               [<AxesSubplot:xlabel='Id', ylabel='PetalWidthCm'>,
               <AxesSubplot:xlabel='SepalLengthCm', ylabel='PetalWidthCm'>,
               <AxesSubplot:xlabel='SepalWidthCm', ylabel='PetalWidthCm'>,
               <AxesSubplot:xlabel='PetalLengthCm', ylabel='PetalWidthCm'>,
               <AxesSubplot:xlabel='PetalWidthCm', ylabel='PetalWidthCm'>]],
dtype=object)

```



```
In [ ]: # 5- Use sns.pairplot to plot data
sns.pairplot(iris_data)
```

```
Out[ ]: <seaborn.axisgrid.PairGrid at 0x1dc06e13640>
```



Congratualtions! You finished your first lab!