importing libraries

```
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import os
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```
import seaborn as sns
import matplotlib.pyplot as plt
import plotly.express as px
import warnings
warnings.filterwarnings('ignore')
```

reading data

```
df1 = pd.read_csv('/content/drive/MyDrive/mental fitness/mental-disease.csv')
df2 =pd.read_csv('/content/drive/MyDrive/mental fitness/prevalence-by-mental-and-substance-use-disorder.csv')
```

```
df1.head()
```

```
        Entity Code  Year  \
0  Afghanistan  AFG  1990
1  Afghanistan  AFG  1991
2  Afghanistan  AFG  1992
3  Afghanistan  AFG  1993
4  Afghanistan  AFG  1994

   DALYs (Disability-Adjusted Life Years) - Mental disorders - Sex: 
Both - Age: All Ages (Percent)
0                                           1.696670

1                                           1.734281

2                                           1.791189

3                                           1.776779

4                                           1.712986
```

```
df2.head()
```

```
        Entity Code  Year  \
0  Afghanistan  AFG  1990
1  Afghanistan  AFG  1991
2  Afghanistan  AFG  1992
3  Afghanistan  AFG  1993
4  Afghanistan  AFG  1994

   Prevalence - Schizophrenia - Sex: Both - Age: Age-standardized
(Percent)  \
0                                           0.228979

1                                           0.228120

2                                           0.227328

3                                           0.226468

4                                           0.225567


   Prevalence - Bipolar disorder - Sex: Both - Age: Age-standardized
(Percent)  \
0                                           0.721207

1                                           0.719952

2                                           0.718418

3                                           0.717452

4                                           0.717012


   Prevalence - Eating disorders - Sex: Both - Age: Age-standardized
(Percent)  \
0                                           0.131001

1                                           0.126395

2                                           0.121832

3                                           0.117942

4                                           0.114547


   Prevalence - Anxiety disorders - Sex: Both - Age: Age-standardized
(Percent)  \
0                                           4.835127

1                                           4.821765
```

```
2                                           4.801434

3                                           4.789363

4                                           4.784923


   Prevalence - Drug use disorders - Sex: Both - Age: Age-standardized
(Percent)  \
0                                           0.454202

1                                           0.447112

2                                           0.441190

3                                           0.435581

4                                           0.431822


   Prevalence - Depressive disorders - Sex: Both - Age: Age-
standardized (Percent)  \
0                                           5.125291

1                                           5.116306

2                                           5.106558

3                                           5.100328

4                                           5.099424


   Prevalence - Alcohol use disorders - Sex: Both - Age: Age-
standardized (Percent)
0                                           0.444036

1                                           0.444250

2                                           0.445501

3                                           0.445958

4                                           0.445779
```

merging two datasets

```
data = pd.merge(df1, df2)
data.head()
```

```
        Entity Code  Year  \
0  Afghanistan  AFG  1990
1  Afghanistan  AFG  1991
2  Afghanistan  AFG  1992
3  Afghanistan  AFG  1993
4  Afghanistan  AFG  1994

   DALYs (Disability-Adjusted Life Years) - Mental disorders - Sex:
Both - Age: All Ages (Percent)  \
0                                            1.696670

1                                            1.734281

2                                            1.791189

3                                            1.776779

4                                            1.712986


   Prevalence - Schizophrenia - Sex: Both - Age: Age-standardized
(Percent)  \
0                                            0.228979

1                                            0.228120

2                                            0.227328

3                                            0.226468

4                                            0.225567


   Prevalence - Bipolar disorder - Sex: Both - Age: Age-standardized
(Percent)  \
0                                            0.721207

1                                            0.719952

2                                            0.718418

3                                            0.717452

4                                            0.717012


   Prevalence - Eating disorders - Sex: Both - Age: Age-standardized
(Percent)  \
0                                            0.131001

1                                            0.126395
```

```
2                                              0.121832

3                                              0.117942

4                                              0.114547


   Prevalence - Anxiety disorders - Sex: Both - Age: Age-standardized
(Percent)  \
0                                              4.835127

1                                              4.821765

2                                              4.801434

3                                              4.789363

4                                              4.784923


   Prevalence - Drug use disorders - Sex: Both - Age: Age-standardized
(Percent)  \
0                                              0.454202

1                                              0.447112

2                                              0.441190

3                                              0.435581

4                                              0.431822


   Prevalence - Depressive disorders - Sex: Both - Age: Age-
standardized (Percent)  \
0                                              5.125291

1                                              5.116306

2                                              5.106558

3                                              5.100328

4                                              5.099424


   Prevalence - Alcohol use disorders - Sex: Both - Age: Age-
standardized (Percent)
0                                              0.444036
```

```
1                                               0.444250

2                                               0.445501

3                                               0.445958

4                                               0.445779
```

data cleaning

```
data.isnull().sum()
```

```
Entity
0
Code
690
Year
0
DALYs (Disability-Adjusted Life Years) - Mental disorders - Sex: Both
- Age: All Ages (Percent)      0
Prevalence - Schizophrenia - Sex: Both - Age: Age-standardized
(Percent)                                 0
Prevalence - Bipolar disorder - Sex: Both - Age: Age-standardized
(Percent)                              0
Prevalence - Eating disorders - Sex: Both - Age: Age-standardized
(Percent)                              0
Prevalence - Anxiety disorders - Sex: Both - Age: Age-standardized
(Percent)                             0
Prevalence - Drug use disorders - Sex: Both - Age: Age-standardized
(Percent)                            0
Prevalence - Depressive disorders - Sex: Both - Age: Age-standardized
(Percent)                         0
Prevalence - Alcohol use disorders - Sex: Both - Age: Age-standardized
(Percent)                       0
dtype: int64
```

```
data.drop('Code',axis=1,inplace=True)
```

```
data.head()
```

```
        Entity  Year  \
0  Afghanistan  1990
1  Afghanistan  1991
2  Afghanistan  1992
3  Afghanistan  1993
4  Afghanistan  1994

   DALYs (Disability-Adjusted Life Years) - Mental disorders - Sex:
Both - Age: All Ages (Percent)  \
0                                                 1.696670

1                                                 1.734281

2                                                 1.791189

3                                                 1.776779

4                                                 1.712986


   Prevalence - Schizophrenia - Sex: Both - Age: Age-standardized
(Percent)  \
0                                                 0.228979

1                                                 0.228120

2                                                 0.227328

3                                                 0.226468

4                                                 0.225567


   Prevalence - Bipolar disorder - Sex: Both - Age: Age-standardized
(Percent)  \
0                                                 0.721207

1                                                 0.719952

2                                                 0.718418

3                                                 0.717452

4                                                 0.717012


   Prevalence - Eating disorders - Sex: Both - Age: Age-standardized
(Percent)  \
0                                                 0.131001

1                                                 0.126395

2                                                 0.121832

3                                                 0.117942

4                                                 0.114547


   Prevalence - Anxiety disorders - Sex: Both - Age: Age-standardized
```

```
(Percent)  \
0                                                4.835127

1                                                4.821765

2                                                4.801434

3                                                4.789363

4                                                4.784923


   Prevalence - Drug use disorders - Sex: Both - Age: Age-standardized
(Percent)  \
0                                                0.454202

1                                                0.447112

2                                                0.441190

3                                                0.435581

4                                                0.431822


   Prevalence - Depressive disorders - Sex: Both - Age: Age-
standardized (Percent)  \
0                                                5.125291

1                                                5.116306

2                                                5.106558

3                                                5.100328

4                                                5.099424


   Prevalence - Alcohol use disorders - Sex: Both - Age: Age-
standardized (Percent)
0                                                0.444036

1                                                0.444250

2                                                0.445501

3                                                0.445958

4                                                0.445779

data.size,data.shape
```

```
(68400, (6840, 10))
```

```
data.set_axis(['Country','Year','Schizophrenia', 'Bipolar_disorder',
'Eating_disorder','Anxiety','drug_usage','depression','alcohol','menta
l_fitness'], axis='columns', inplace=True)
```

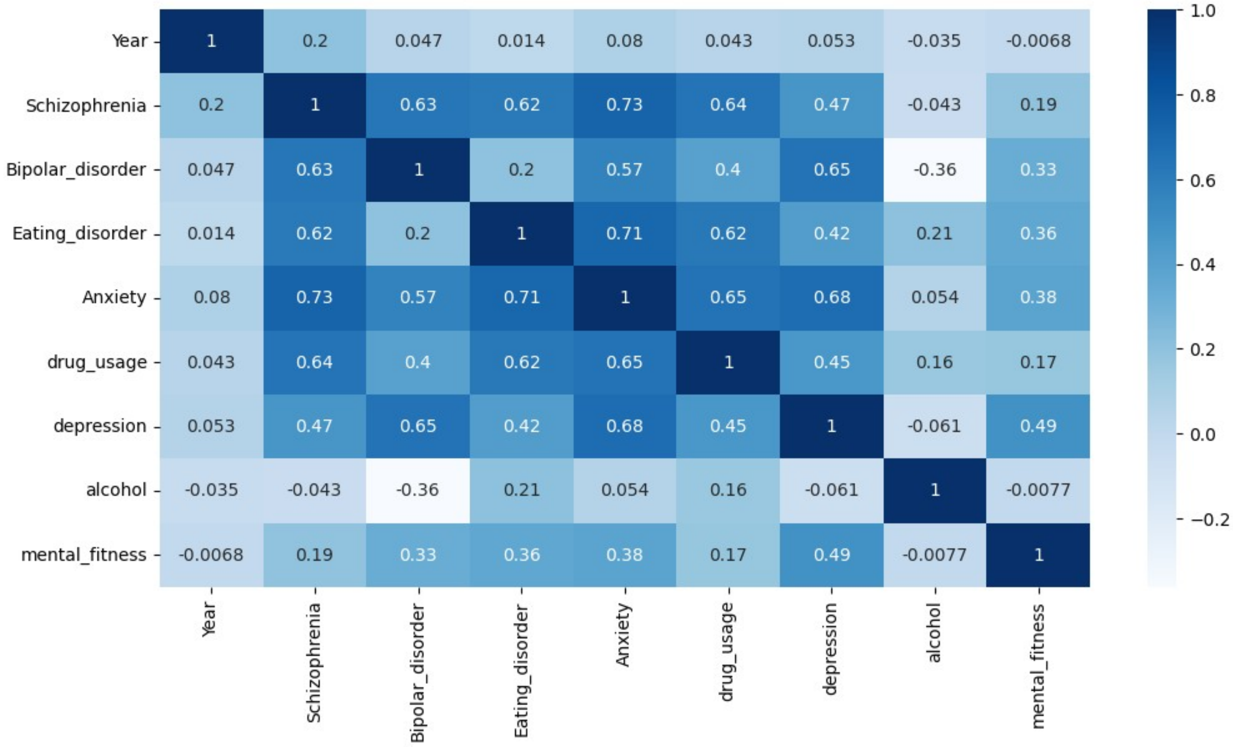
```
data.head()
```

```
       Country  Year  Schizophrenia  Bipolar_disorder  Eating_disorder
\
0  Afghanistan  1990       1.696670          0.228979         0.721207

1  Afghanistan  1991       1.734281          0.228120         0.719952

2  Afghanistan  1992       1.791189          0.227328         0.718418

3  Afghanistan  1993       1.776779          0.226468         0.717452

4  Afghanistan  1994       1.712986          0.225567         0.717012


    Anxiety  drug_usage  depression   alcohol  mental_fitness
0  0.131001    4.835127    0.454202  5.125291        0.444036
1  0.126395    4.821765    0.447112  5.116306        0.444250
2  0.121832    4.801434    0.441190  5.106558        0.445501
3  0.117942    4.789363    0.435581  5.100328        0.445958
4  0.114547    4.784923    0.431822  5.099424        0.445779
```

explaratory data analysis(EDA)

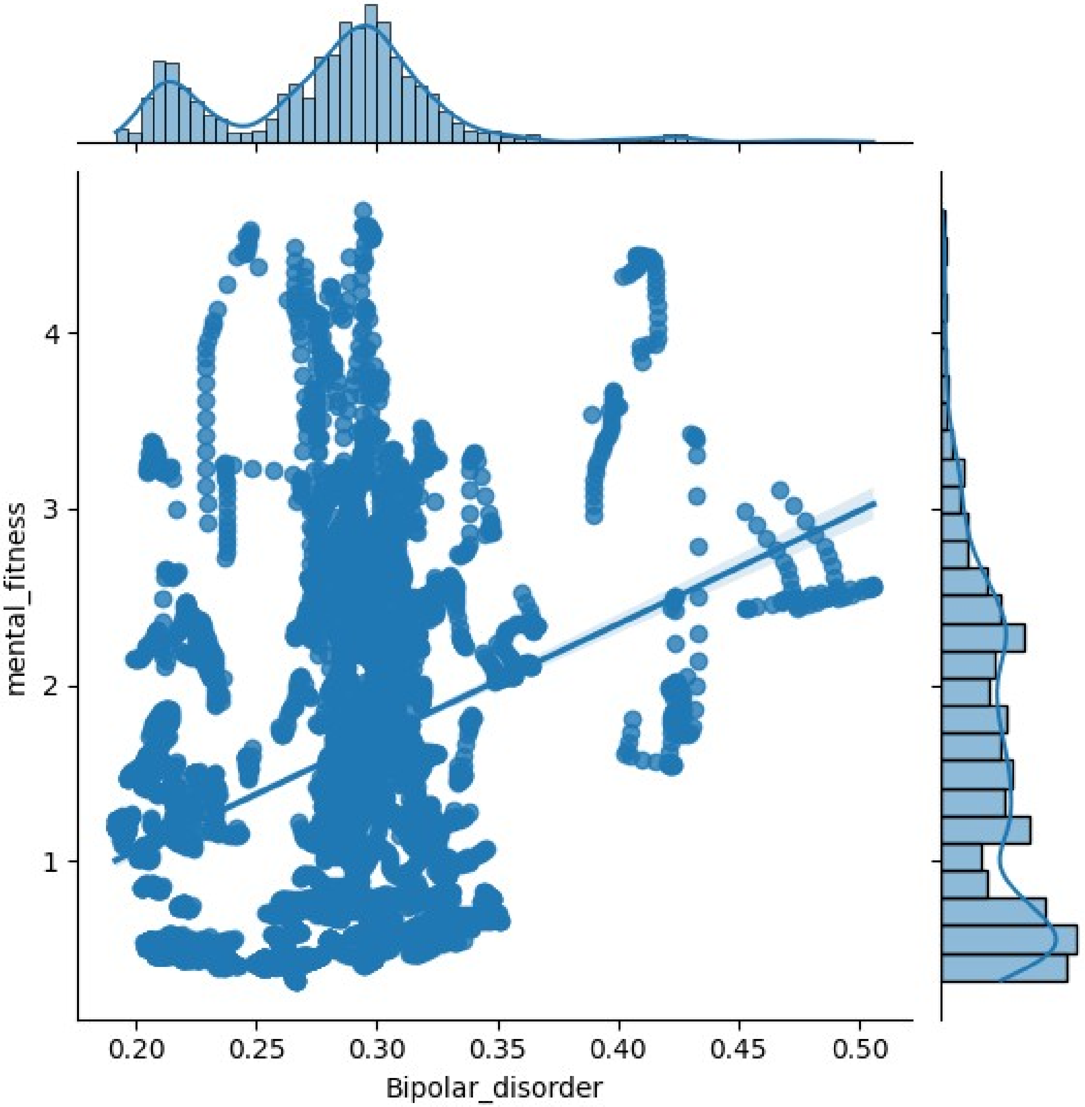
```
plt.figure(figsize=(12,6))
sns.heatmap(data.corr(),annot=True,cmap='Blues')
plt.plot()
```

```
[]
```

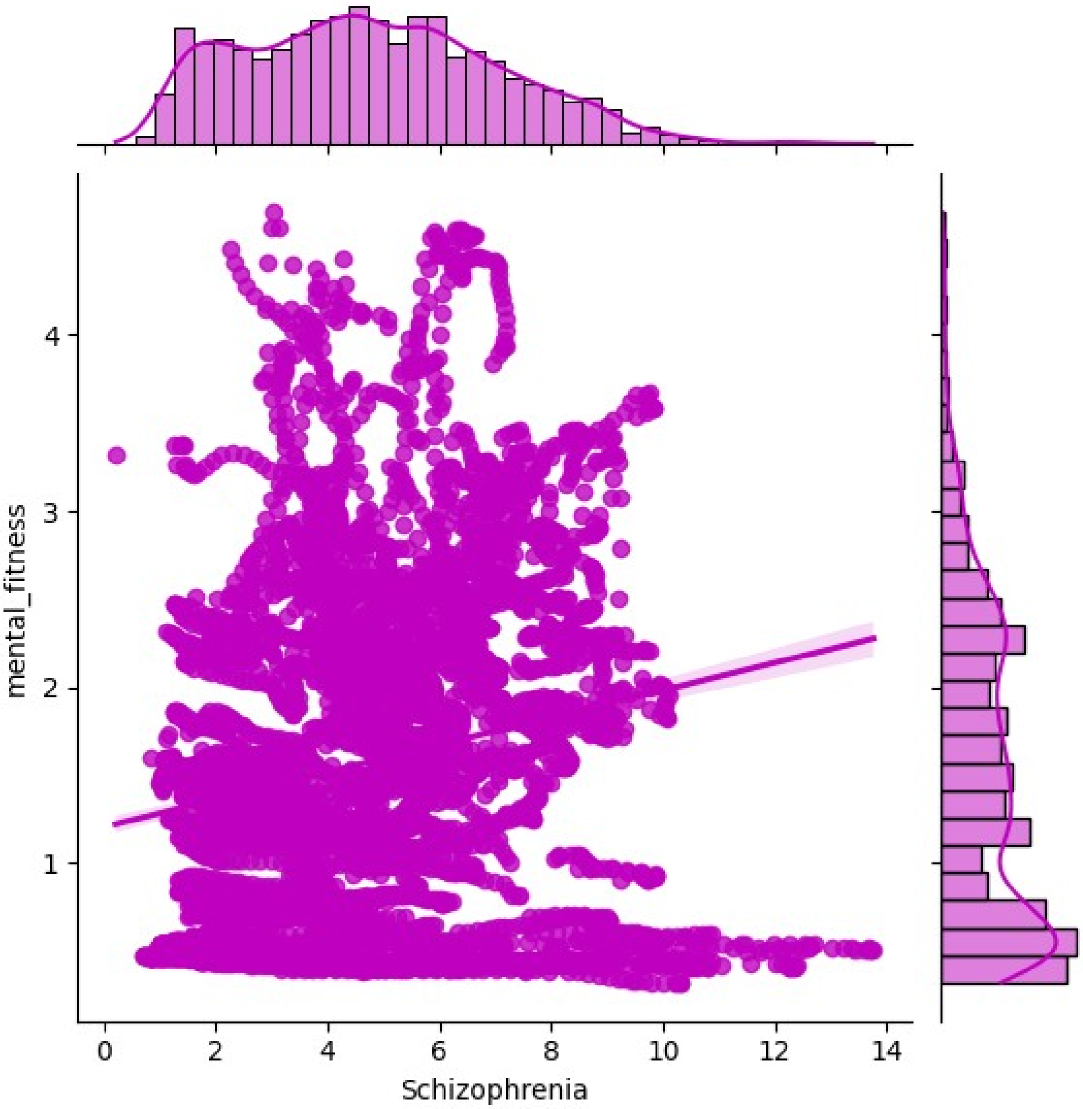


	Year	Schizophrenia	Bipolar_disorder	Eating_disorder	Anxiety	drug_usage	depression	alcohol	mental_fitness
Year	1	0.2	0.047	0.014	0.08	0.043	0.053	-0.035	-0.0068
Schizophrenia	0.2	1	0.63	0.62	0.73	0.64	0.47	-0.043	0.19
Bipolar_disorder	0.047	0.63	1	0.2	0.57	0.4	0.65	-0.36	0.33
Eating_disorder	0.014	0.62	0.2	1	0.71	0.62	0.42	0.21	0.36
Anxiety	0.08	0.73	0.57	0.71	1	0.65	0.68	0.054	0.38
drug_usage	0.043	0.64	0.4	0.62	0.65	1	0.45	0.16	0.17
depression	0.053	0.47	0.65	0.42	0.68	0.45	1	-0.061	0.49
alcohol	-0.035	-0.043	-0.36	0.21	0.054	0.16	-0.061	1	-0.0077
mental_fitness	-0.0068	0.19	0.33	0.36	0.38	0.17	0.49	-0.0077	1

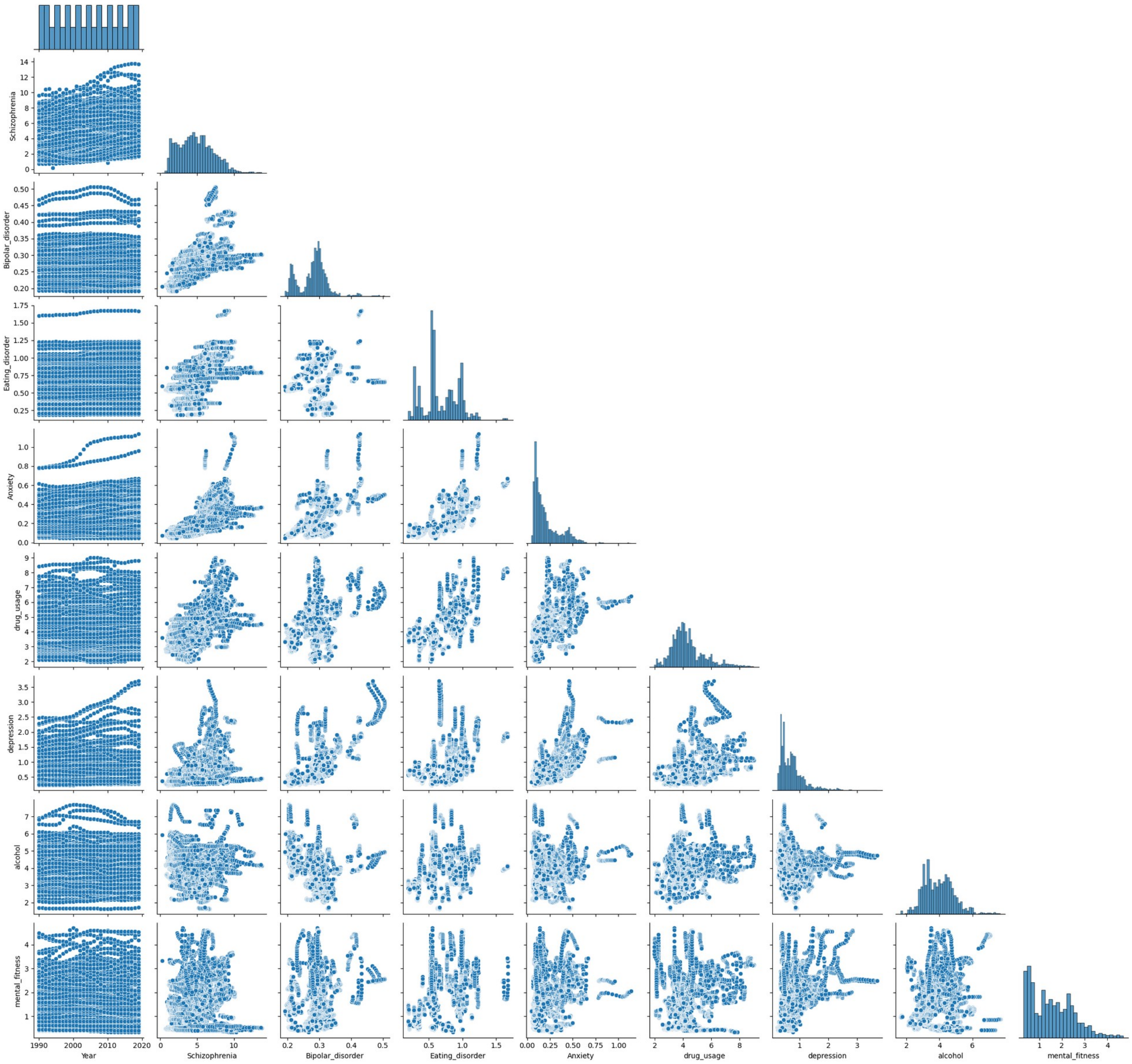
```
sns.jointplot(data=data,x='Bipolar_disorder',y='mental_fitness',kind='
reg')
plt.show()
```



```
sns.jointplot(x='Schizophrenia',y='mental_fitness',data=data,kind='reg
',color='m')
plt.show()
```



```
sns.pairplot(data,corner=True)
plt.show()
```



```
mean = data['mental_fitness'].mean()
mean
```

```
1.5788071625377194
```

```
fig = px.pie(data, values='mental_fitness', names='Year')
fig.show()
```

year wise variation in mental fitness in different countries

```
fig = px.line(data, x="Year", y="mental_fitness",
color='Country',markers=True,color_discrete_sequence=['red','blue'],te
mplate='plotly_dark')
fig.show()
```

```
df = data.copy()
```

```
df.head()
```

```
       Country  Year  Schizophrenia  Bipolar_disorder  Eating_disorder  
\
0  Afghanistan  1990       1.696670          0.228979         0.721207   

1  Afghanistan  1991       1.734281          0.228120         0.719952   

2  Afghanistan  1992       1.791189          0.227328         0.718418   

3  Afghanistan  1993       1.776779          0.226468         0.717452   

4  Afghanistan  1994       1.712986          0.225567         0.717012   


    Anxiety  drug_usage  depression   alcohol  mental_fitness  
0  0.131001    4.835127    0.454202  5.125291        0.444036  
1  0.126395    4.821765    0.447112  5.116306        0.444250  
2  0.121832    4.801434    0.441190  5.106558        0.445501  
3  0.117942    4.789363    0.435581  5.100328        0.445958  
4  0.114547    4.784923    0.431822  5.099424        0.445779  
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 6840 entries, 0 to 6839
Data columns (total 10 columns):
 #   Column            Non-Null Count  Dtype  
---  ------            --------------  -----  
 0   Country           6840 non-null   object 
 1   Year              6840 non-null   int64  
 2   Schizophrenia     6840 non-null   float64
 3   Bipolar_disorder  6840 non-null   float64
 4   Eating_disorder   6840 non-null   float64
 5   Anxiety           6840 non-null   float64
 6   drug_usage        6840 non-null   float64
 7   depression        6840 non-null   float64
 8   alcohol           6840 non-null   float64
 9   mental_fitness    6840 non-null   float64
```

```
dtypes: float64(8), int64(1), object(1)
memory usage: 587.8+ KB

from sklearn.preprocessing import LabelEncoder
l=LabelEncoder()
for i in df.columns:
    if df[i].dtype == 'object':
        df[i]=l.fit_transform(df[i])

X = df.drop('mental_fitness',axis=1)
y = df['mental_fitness']

from sklearn.model_selection import train_test_split
xtrain, xtest, ytrain, ytest = train_test_split(X, y, test_size=0.2,
random_state=2)

X = df.drop('mental_fitness',axis=1)
y = df['mental_fitness']

from sklearn.model_selection import train_test_split
xtrain, xtest, ytrain, ytest = train_test_split(X, y, test_size=0.2,
random_state=2)
```

applying classification algorithm

linear regression

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
lr = LinearRegression()
lr.fit(xtrain,ytrain)

# model evaluation for training set
ytrain_pred = lr.predict(xtrain)
mse = mean_squared_error(ytrain, ytrain_pred)
rmse = (np.sqrt(mean_squared_error(ytrain, ytrain_pred)))
r2 = r2_score(ytrain, ytrain_pred)

print("The model performance for training set")
print("--------------------------------------")
print('MSE is {}'.format(mse))
print('RMSE is {}'.format(rmse))
print('R2 score is {}'.format(r2))
print("\n")

# model evaluation for testing set
ytest_pred = lr.predict(xtest)
```

```
mse = mean_squared_error(ytest, ytest_pred)
rmse = (np.sqrt(mean_squared_error(ytest, ytest_pred)))
r2 = r2_score(ytest, ytest_pred)

print("The model performance for testing set")
print("--------------------------------------")
print('MSE is {}'.format(mse))
print('RMSE is {}'.format(rmse))
print('R2 score is {}'.format(r2))
```

```
The model performance for training set
--------------------------------------
MSE is 0.576867540006079
RMSE is 0.7595179655584712
R2 score is 0.33581211668682887


The model performance for testing set
--------------------------------------
MSE is 0.5792230514362919
RMSE is 0.7610670479243546
R2 score is 0.3513086902731888
```

random forest regression

```
from sklearn.ensemble import RandomForestRegressor
rf = RandomForestRegressor()
rf.fit(xtrain, ytrain)

# model evaluation for training set
ytrain_pred = rf.predict(xtrain)
mse = mean_squared_error(ytrain, ytrain_pred)
rmse = (np.sqrt(mean_squared_error(ytrain, ytrain_pred)))
r2 = r2_score(ytrain, ytrain_pred)

print("The model performance for training set")
print("--------------------------------------")
print('MSE is {}'.format(mse))
print('RMSE is {}'.format(rmse))
print('R2 score is {}'.format(r2))
print("\n")

# model evaluation for testing set
ytest_pred = rf.predict(xtest)
mse = mean_squared_error(ytest, ytest_pred)
rmse = (np.sqrt(mean_squared_error(ytest, ytest_pred)))
r2 = r2_score(ytest, ytest_pred)
```

```
print("The model performance for testing set")
print("--------------------------------------")
print('MSE is {}'.format(mse))
print('RMSE is {}'.format(rmse))
print('R2 score is {}'.format(r2))
```

```
The model performance for training set
--------------------------------------
MSE is 0.0006193541437514021
RMSE is 0.02488682671116191
R2 score is 0.9992868943228195


The model performance for testing set
--------------------------------------
MSE is 0.0036496873709222374
RMSE is 0.06041264247591093
R2 score is 0.9959125927829249
```

decission tree regression

```
from sklearn.tree import DecisionTreeRegressor
dtr=DecisionTreeRegressor(random_state=0)
dtr.fit(xtrain,ytrain)

#predicting the value

ytest_pred=dtr.predict(xtest)

#evaluating the model

from sklearn.metrics import r2_score,mean_squared_error
print("Results for Decision Tree Regression:\n1)Mean Square Error={}\n2)R-Square Score={}".format(mean_squared_error(ytest,ytest_pred),r2_score(ytest,ytest_pred)))
```

```
Results for Decision Tree Regression:
1)Mean Square Error=0.012478922917737525
2)R-Square Score=0.9860244359553473
```

svm regression

```
from sklearn.svm import SVR
svr=SVR()
svr.fit(xtrain,ytrain)
```

```
#predicting the value

ytest_pred=svr.predict(xtest)

#evaluating the model

from sklearn.metrics import r2_score,mean_squared_error
print("Results for SVM Regression:\n1)Mean Square Error={}\n2)R-Square
Score={}".format(mean_squared_error(ytest,ytest_pred),r2_score(ytest,y
test_pred)))
```

```
Results for SVM Regression:
1)Mean Square Error=0.902874531063811
2)R-Square Score=-0.011159449925972176
```

conclusion

Random Forest Regression works well on both train and test sets with r2 score of 0.99. As well as Decision Tree Regression also works well on both train and test set with r2 score of 0.98.

```
np.random.seed(range(0,100))
print("Welcome to Mental Fitness Tracker!\nFill the detail to check
your mental fitness!")
country=l.fit_transform([input('Enter Your country Name:')])
year=int(input("Enter the Year:"))
schi=(float(input("Enter your Schizophrenia rate in % (it not enter
0):")))*100
bipo_dis=(float(input("Enter your Bipolar disorder rate in % (it not
enter 0):")))*100
eat_dis=(float(input("Enter your Eating disorder rate in % (it not
enter 0):")))*100
anx=(float(input("Enter your Anxiety rate in % (it not enter
0):")))*100
drug_use=(float(input("Enter your Drug Usage rate in per year % (it
not enter 0):")))*100
depr=(float(input("Enter your Depression rate in % (it not enter
0):")))*100
alch=(float(input("Enter your Alcohol Consuming rate per year in % (it
not enter 0):")))*100

prediction=rf.predict([[country,year,schi,bipo_dis,eat_dis,anx,drug_us
e,depr,alch]])
print("Your Mental Fitness is {}%".format(prediction))
print("Bye...!")
```

```
Welcome to Mental Fitness Tracker!
Fill the detail to check your mental fitness!
Enter Your country Name:pakistan
Enter the Year:87
Enter your Schizophrenia rate in % (it not enter 0):98
Enter your Bipolar disorder rate in % (it not enter 0):78
Enter your Eating disorder rate in % (it not enter 0):76
Enter your Anxiety rate in % (it not enter 0):99
Enter your Drug Usage rate in per year % (it not enter 0):78
Enter your Depression rate in % (it not enter 0):85
Enter your Alcohol Consuming rate per year in % (it not enter 0):67
Your Mental Fitness is [4.02270229]%
Bye...!
```