

Meet - anz-inoj-mdk

Mail - R GOUTAM - Outlook

190031925-ds lab 4 .ipynb - Cola

colab.research.google.com/drive/1AxIOLc4OiUqdbZig4nqFqhtZNFdh3ol#scrollTo=6swarWoO4eMO

Apps190031925 Mail - R...KL-ONLINE-CLASSCoursera | kd-2020-...Global NetAcad Ins...Wiley Certification I...Projects · Dashboar...Online Python Com...

190031925-ds lab 4 .ipynb

FileEditViewInsertRuntimeToolsHelpAll changes saved

Files

sample_data

2015 (1).csv

2015.csv

2016 (1).csv

2016.csv

2017 (1).csv

2017.csv

2018 (1).csv

2018.csv

2019 (1).csv

2019.csv

graph.png

190031925-r.goutam-ds lab 4

import numpy as np

import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

import warnings

warnings.filterwarnings("ignore")

from google.colab import files

uploaded=files.upload()

Choose Files5 files

2015.csv(application/vnd.ms-excel) - 16557 bytes, last modified: 1/27/2021 - 100% done

2016.csv(application/vnd.ms-excel) - 17132 bytes, last modified: 1/27/2021 - 100% done

2017.csv(application/vnd.ms-excel) - 29536 bytes, last modified: 1/27/2021 - 100% done

2018.csv(application/vnd.ms-excel) - 8809 bytes, last modified: 1/27/2021 - 100% done

2019.csv(application/vnd.ms-excel) - 8822 bytes, last modified: 1/27/2021 - 100% done

Saving 2015.csv to 2015 (1).csv

Saving 2016.csv to 2016 (1).csv

Saving 2017.csv to 2017 (1).csv

Saving 2018.csv to 2018 (1).csv

Saving 2019.csv to 2019 (1).csv

[3] data_2015=pd.read_csv('2015.csv')

data_2016=pd.read_csv('2016.csv')

data_2017=pd.read_csv('2017.csv')

data_2015.columns=[each.split()[0] if(len(each.split())>2) else each.replace(" ","_") for each in data_2015.columns]

data_2016.columns=[each.split()[0] if(len(each.split())>2) else each.replace(" ","_") for each in data_2016.columns]

data_2017.columns=[each.replace(" ","_") for each in data_2017.columns]


data_2017.columns=[each.split()[0] if(len(each.split())>2) else each.replace(" ","_") for each in data_2017.columns]

data_2015.head()

	Country	Region	Happiness_Rank	Happiness_Score	Standard_Error	Economy	Family	Health	Freedom	Trust	Generosity	Dystopia_Residual
0	Switzerland	Western Europe	1	7.587	0.03411	1.39651	1.34951	0.94143	0.65557	0.41978	0.29678	2.51738
1	Iceland	Western Europe	2	7.561	0.04884	1.30232	1.40223	0.94784	0.62877	0.14145	0.43630	2.70201
2	Denmark	Western Europe	3	7.527	0.03328	1.32548	1.36058	0.87464	0.64938	0.48357	0.34139	2.49204

Disk75.75 GB available

Person 1



Goutam

goutamrayavarapugoutam@gmail.com

Sync is off

Turn on sync...

Manage your Google Account

Sign out

Other people

goutam

Goutam

Guest

Add

Meet - anz-inoj-mdk

Mail - R GOUTAM - Outlook

190031925-ds lab 4 .ipynb - Cola

colab.research.google.com/drive/1AxLOLc4OiUqdbZig4nqFfghtZNFdh3ol#scrollTo=PMkh-XR35afx

Apps190031925 Mail - R...KL-ONLINE-CLASSCoursera | kd-2020-...Global NetAcad Ins...Wiley Certification I...Projects · Dashboard...Online Python Com...

190031925-ds lab 4 .ipynb

FileEditViewInsertRuntimeToolsHelpAll changes saved

Files

sample_data2015 (1).csv2015.csv2016 (1).csv2016.csv2017 (1).csv2017.csv2018 (1).csv2018.csv2019 (1).csv2019.csvgraph.png

+ Code+ Text

[3] 3 Norway Western Europe 4 7.522 0.03680 1.45900 1.33095 0.88521 0.66973 0.36503 0.34699
4 Canada North America 5 7.427 0.03553 1.32629 1.32261 0.90563 0.63297 0.32957 0.45611

print("Are There Missing Data? :",data_2015.isnull().any().any())
print(data_2015.isnull().sum())

Are There Missing Data? : False
Country 0
Region 0
Happiness_Rank 0
Happiness_Score 0
Standard_Error 0
Economy 0
Family 0
Health 0
Freedom 0
Trust 0
Generosity 0
Dystopia_Residual 0
dtype: int64


[5] region_lists=list(data_2015['Region'].unique())
region_happiness_ratio=[]
for each in region_lists:
region=data_2015[data_2015['Region']==each]
region_happiness_rate=sum(region.Happiness_Score)/len(region)
region_happiness_ratio.append(region_happiness_rate)

data=pd.DataFrame({'region':region_lists,'region_happiness_ratio':region_happiness_ratio})
new_index=(data['region_happiness_ratio'].sort_values(ascending=False)).index.values
sorted_data =data.reindex(new_index)




sorted_data

	region	region_happiness_ratio
2	Australia and New Zealand	7.285000
1	North America	7.273000
0	Western Europe	6.689619
4	Latin America and Caribbean	6.144682
7	Eastern Asia	5.626167

Person 1



Goutam
goutamrayavarapugoutam@gmail.com



Sync is off

Turn on sync...

Manage your Google Account

Sign out

Other people

goutam

Goutam

Guest

Add

Disk75.75 GB available

190031925-ds lab 4 .ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

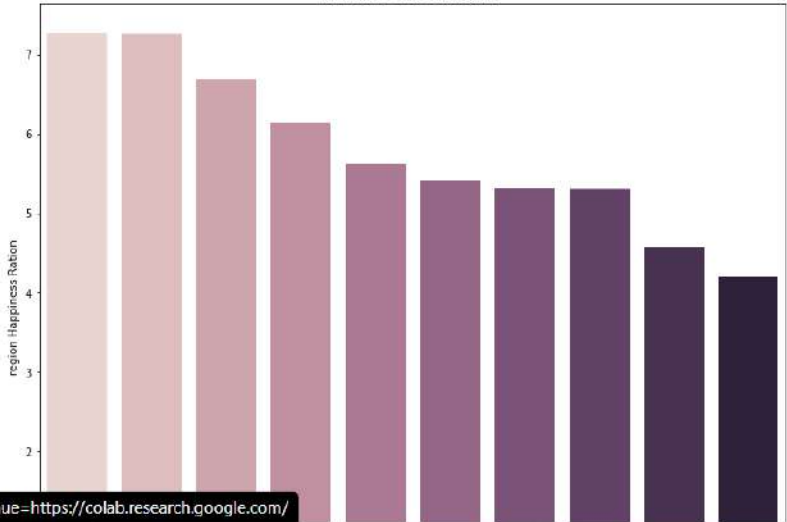
- sample_data
- 2015 (1).csv
- 2015.csv
- 2016 (1).csv
- 2016.csv
- 2017 (1).csv
- 2017.csv
- 2018 (1).csv
- 2018.csv
- 2019 (1).csv
- 2019.csv
- graph.png

+ Code + Text

3	Middle East and Northern Africa	5.406900
6	Central and Eastern Europe	5.332931
5	Southeastern Asia	5.317444
9	Southern Asia	4.580857
8	Sub-Saharan Africa	4.202800


```
[6] #barplot
plt.figure(figsize=(12,10))
sns.barplot(x=sorted_data['region'], y=sorted_data['region_happiness_ratio'], palette=sns.cubehelix_palette(len(sorted_data['region'])))
plt.xticks(rotation = 90)
plt.xlabel('region')
plt.ylabel('region Happiness Ratio')
plt.title('Happiness rate for regions')
plt.show()
```


Happiness rate for regions





region	region_happiness_ratio
Middle East and Northern Africa	5.406900
Central and Eastern Europe	5.332931
Southeastern Asia	5.317444
Southern Asia	4.580857
Sub-Saharan Africa	4.202800


Comment Share


Goutam
goutamrayavarepugoutam@gmail.com
Manage your Google Account

 priyanka reddy
priyankareddy306102@gmail.com

 goutam 29
g9110793335@gmail.com

 goutamr2915@gmail.com
goutamr2915@gmail.com

 goutam r
rayavarapugoutam29@gmail.com

+ Add another account

Sign out of all accounts

Privacy Policy · Terms of Service

https://accounts.google.com/SignOutOptions?hl=en&continue=https://colab.research.google.com/

Meet - anz-inoj-mdk

Mail - R GOUTAM - Outlook

190031925-ds lab 4 .ipynb - Cola

colab.research.google.com/drive/1AxlOLc4OiUqdbZig4nqFfghtZNFdh3ol#scrollTo=aZdZUm9k5c0v

Apps190031925 Mail - R...KL-ONLINE-CLASSCoursera | kd-2020-...Global NetAcad Ins...Wiley Certification I...Projects · Dashboard...Online Python Com...

190031925-ds lab 4 .ipynb

FileEditViewInsertRuntimeToolsHelpAll changes saved

Files

sample_data

2015 (1).csv

2015.csv

2016 (1).csv

2016.csv

2017 (1).csv

2017.csv

2018 (1).csv

2018.csv

2019 (1).csv

2019.csv

graph.png

```
[7] region_lists=list(data_2015['Region'].unique())
region_economy_ratio=[]
for each in region_lists:
    region=data_2015[data_2015['Region']==each]
    region_economy_rate=sum(region.Economy)/len(region)
    region_economy_ratio.append(region_economy_rate)

data_economy=pd.DataFrame({'region':region_lists,'region_economy_ratio':region_economy_ratio})
new_index_economy=(data_economy['region_economy_ratio'].sort_values(ascending=True)).index.values
sorted_data_economy =data_economy.reindex(new_index)

sorted_data_economy
```


	region	region_economy_ratio
2	Australia and New Zealand	1.291880
1	North America	1.360400
0	Western Europe	1.298596
4	Latin America and Caribbean	0.876815
7	Eastern Asia	1.151780
3	Middle East and Northern Africa	1.066973
6	Central and Eastern Europe	0.942438
5	Southeastern Asia	0.789054
9	Southern Asia	0.560486
8	Sub-Saharan Africa	0.380473

```
[8] f,ax1=plt.subplots(figsize=(12,10))
sns.barplot(x=sorted_data_economy['region'],y=sorted_data_economy['region_economy_ratio'],palette="rocket",ax=ax1)
plt.xticks(rotation=90)
plt.xlabel('Region')
plt.ylabel('Region Economy Ratio')
plt.title('Economy rate for regions')
plt.show()
```

Economy rate for regions




Disk75.75 GB available


Person 1




Goutam


goutamrayavarapugoutam@gmail.com





 Sync is off


Turn on sync...


 Manage your Google Account

 Sign out

Other people

 9 goutam

 Goutam

 Guest

+ Add

Meet - anz-inoj-mdk

Mail - R GOUTAM - Outlook

190031925-ds lab 4 .ipynb - Cola

colab.research.google.com/drive/1AxIOLc4OiUqdbZig4nqFfqhtZNFdh3ol#scrollTo=aZdZUmkg5c0v

Apps190031925 Mail - R...KL-ONLINE-CLASSCoursera | kd-2020-...Global NetAcad Ins...Wiley Certification I...Projects · Dashboar...Online Python Com...

190031925-ds lab 4 .ipynb

FileEditViewInsertRuntimeToolsHelpAll changes saved

Files

sample_data

2015 (1).csv

2015.csv

2016 (1).csv

2016.csv

2017 (1).csv

2017.csv

2018 (1).csv

2018.csv

2019 (1).csv

2019.csv

graph.png

[8]

Economy rate for regions

Region	Region Economy Ratio
Australia and New Zealand	1.3
North America	1.35
Western Europe	1.3
Latin America and Caribbean	0.85
Eastern Asia	1.15
Middle East and Northern Africa	1.05
Central and Eastern Europe	0.95
Southeastern Asia	0.8
Southern Asia	0.55
Sub-Saharan Africa	0.38

[9]

#Point plot
sorted_data['region_happiness_ratio']-sorted_data['region_happiness_ratio']/max(sorted_data['region_happiness_ratio'])

Disk75.75 GB available

Person 1

Goutam

goutamrayavarapugoutam@gmail.com

Sync is off

Turn on sync...

Manage your Google Account

Sign out

Other people

goutam

Goutam

Guest

Add

Meet - anz-inoj-mdk

Mail - R GOUTAM - Outlook

190031925-ds lab 4 .ipynb - Cola

colab.research.google.com/drive/1AxIOLc4OiUqdbZig4nqFfghtZNFdh3oI#scrollTo=aZdZUm9k5c0v

Apps190031925 Mail - R...KL-ONLINE-CLASSCoursera | kd-2020-...Global NetAcad Ins...Wiley Certification I...Projects · Dashboar...Online Python Com...

190031925-ds lab 4 .ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_data2015 (1).csv2015.csv2016 (1).csv2016.csv2017 (1).CSV2017.csv2018 (1).csv2018.csv2019 (1).csv2019.csvgraph.png

```
[9] #Point plot
sorted_data['region_happiness_ratio']=sorted_data['region_happiness_ratio']/max(sorted_data['region_happiness_ratio'])
sorted_data_economy['region_economy_ratio']=sorted_data_economy['region_economy_ratio']/max(sorted_data_economy['region_economy_ratio'])

data=pd.concat([sorted_data,sorted_data_economy['region_economy_ratio']],axis=1)
data.sort_values('region_happiness_ratio',inplace=True)

f,ax1=plt.subplots(figsize=(20,10))
sns.pointplot(x='region',y='region_happiness_ratio',data=data,color='lime',alpha=0.8)
sns.pointplot(x='region',y='region_economy_ratio',data=data,color='red',alpha=0.8)
plt.text(7.55,0.6,'happiness score ratio',color='red',fontsize=17,style='italic')
plt.text(7.55,0.55,'economy ratio',color='lime',fontsize=18,style='italic')
plt.xticks(rotation=45)
plt.xlabel('Region',fontsize=15,color='blue')
plt.ylabel('Values',fontsize=15,color='blue')
plt.title('Happiness Score VS Economy Rate',fontsize=20,color='blue')
plt.grid()
plt.show()
```

Happiness Score VS Economy Rate

Region	happiness score ratio	economy ratio
...	0.58	0.58
...	0.68	0.68
...	0.78	0.78
...	0.88	0.88
...	0.98	0.98
...	1.0	1.0

Disk75.75 GB available

Person 1

Goutam

goutamrayavarapugoutam@gmail.com

Sync is off

Turn on sync...

Manage your Google Account

Sign out

Other people

Goutam

Goutam

Guest

+ Add

Meet - anz-inoj-mdk

Mail - R GOUTAM - Outlook

190031925-ds lab 4 .ipynb - Cola

colab.research.google.com/drive/1AxLOLc4OiUqdbZig4nqFfghtZNFdh3ol#scrollTo=cFZgsOb-57bT

Apps190031925 Mail - R...KL-ONLINE-CLASSCoursera | kd-2020-...Global NetAcad Ins...Wiley Certification I...Projects · Dashboar...Online Python Com...

190031925-ds lab 4 .ipynb

FileEditViewInsertRuntimeToolsHelpAll changes saved

Files

sample_data2015 (1).csv2015.csv2016 (1).csv2016.csv2017 (1).csv2017.csv2018 (1).csv2018.csv2019 (1).csv2019.csvgraph.png

+ Code + Text

plt.grid()plt.show()

Happiness Score VS Economy Rate

Region	Happiness Score Ratio (Red)	Economy Ratio (Green)
Sub-Saharan Africa	0.28	0.58
Southern Asia	0.42	0.63
Southeastern Asia	0.58	0.73
Central and Eastern Europe	0.70	0.73
Middle East and Northern Africa	0.78	0.74
Eastern Asia	0.85	0.77
Latin America and Caribbean	0.65	0.84
Western Europe	0.95	0.92
North America	1.00	1.00
Australia and New Zealand	0.95	1.00

[10] #jointplot
dataframe=pd.pivot_table(data_2015,index='Region',values=["Happiness_Score","Family"])

Person 1

Goutam
goutamrayavarapugoutam@gmail.com

Sync is off
Turn on sync...

Manage your Google Account
Sign out

Other people
goutam
Goutam
Guest
Add

Meet - anz-inoj-mdk

Mail - R GOUTAM - Outlook

190031925-ds lab 4 .ipynb - Cola

colab.research.google.com/drive/1AxlOLc4OiUqdbZig4nqFfqhtZNFdh3ol#scrollTo=cFZgsOb-57bT

Apps190031925 Mail - R...KL-ONLINE-CLASSCoursera | kd-2020-...Global NetAcad Ins...Wiley Certification I...Projects · Dashboar...Online Python Com...

190031925-ds lab 4 .ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_data2015 (1).csv2015.csv2016 (1).csv2016.csv2017 (1).csv2017.csv2018 (1).csv2018.csv2019 (1).csv2019.csvgraph.png

```
[10] #jointplot
dataframe=pd.pivot_table(data_2015,index='Region',values=["Happiness_Score","Family"])
dataframe["Happiness_Score"]=dataframe["Happiness_Score"]/max(dataframe["Happiness_Score"])
dataframe["Family"]=dataframe["Family"]/max(dataframe["Family"])
sns.jointplot(dataframe.Family,dataframe.Happiness_Score,kind="kde",height=7,space=0)
plt.savefig('graph.png')
plt.show()
```

```
[11] data_2015["Year"]=2015
data_2016["Year"]=2016
data_2017["Year"]=2017
data_concat=pd.concat([data_2015,data_2016,data_2017],axis=0,sort=False)
df=pd.pivot_table(data_concat,index='Year',values="Happiness_Score")
df
```

Disk75.75 GB available

Person 1

Goutam

goutamrayavarapugoutam@gmail.com

Sync is off

Turn on sync...

Manage your Google Account

Sign out

Other people

goutam

Goutam

Guest

+ Add

Meet - anz-inoj-mdk

Mail - R GOUTAM - Outlook

190031925-ds lab 4 .ipynb - Cola

colab.research.google.com/drive/1AxLOLc4OiUqdbZig4nqFfqhtZNFdh3ol#scrollTo=cFZgsOb-57bT

Apps190031925 Mail - R...KL-ONLINE-CLASSCoursera | kd-2020-...Global NetAcad Ins...Wiley Certification I...Projects · Dashboar...Online Python Com...

190031925-ds lab 4 .ipynb

FileEditViewInsertRuntimeToolsHelpAll changes saved

Files

sample_data2015 (1).csv2015.csv2016 (1).csv2016.csv2017 (1).csv2017.csv2018 (1).csv2018.csv2019 (1).csv2019.csvgraph.png

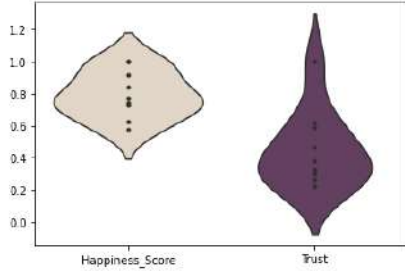
+ Code+ Text

data_2015["Year"]=2015
data_2016["Year"]=2016
data_2017["Year"]=2017
data_concat=pd.concat([data_2015,data_2016,data_2017],axis=0,sort=False)
df=pd.pivot_table(data_concat,index='Year',values="Happiness_Score")
df

Happiness_Score


Year	Happiness_Score
2015	5.375734
2016	5.382185
2017	5.354019

[12] #violinplot
dataframe2=pd.pivot_table(data_2015,index='Region',values=["Happiness_Score","Trust"])
dataframe2["Happiness_Score"]=dataframe2["Happiness_Score"]/max(dataframe2["Happiness_Score"])
dataframe2["Trust"]=dataframe2["Trust"]/max(dataframe2["Trust"])
pal=sns.cubehelix_palette(2,rot=.5,dark=.3)
sns.violinplot(data=dataframe2,palette=pal,inner="points")
plt.show()



[13] #pairplot
sns.pairplot(data)
plt.show()

Person 1



Goutam

goutamrayavarapugoutam@gmail.com

Sync is off

Turn on sync...

Manage your Google Account

Sign out

Other people

goutam

Goutam

Guest

Add

Meet - anz-inoj-mdk

Mail - R GOUTAM - Outlook

190031925-ds lab 4 .ipynb - Cola

colab.research.google.com/drive/1AxlOLc4OiUqdbZig4nqFqhtZNFdh3ol#scrollTo=cFZgsOb-57bT

Apps190031925 Mail - R...KL-ONLINE-CLASSCoursera | kd-2020-...Global NetAcad Ins...Wiley Certification I...Projects · Dashboar...Online Python Com...

190031925-ds lab 4 .ipynb

FileEditViewInsertRuntimeToolsHelpAll changes saved

Files

sample_data

2015 (1).csv

2015.csv

2016 (1).csv

2016.csv

2017 (1).csv

2017.csv

2018 (1).csv

2018.csv

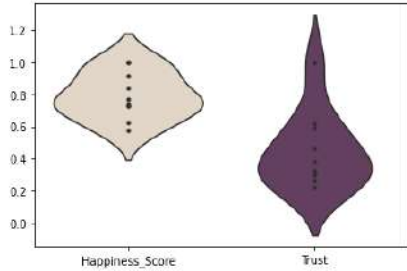
2019 (1).csv

2019.csv

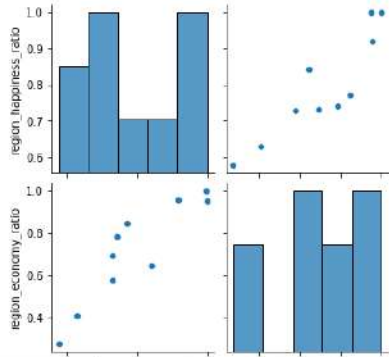
graph.png

+ Code+ Text


```
[12] #violinplot
dataframe2=pd.pivot_table(data_2015,index='Region',values=["Happiness_Score","Trust"])
dataframe2["Happiness_Score"]=dataframe2["Happiness_Score"]/max(dataframe2["Happiness_Score"])
dataframe2["Trust"]=dataframe2["Trust"]/max(dataframe2["Trust"])
pal=sns.cubehelix_palette(2,rot=.5,dark=.3)
sns.violinplot(data=dataframe2,palette=pal,inner="points")
plt.show()
```



```
[13] #pairplot
sns.pairplot(data)
plt.show()
```






Person 1



Goutam

goutamrayavarapugoutam@gmail.com



Sync is off

Turn on sync...

Manage your Google Account

Sign out

Other people

goutam

Goutam

Guest

Add

Disk75.75 GB available

Meet - anz-inoj-mdk

Mail - R GOUTAM - Outlook

190031925-ds lab 4 .ipynb - Cola

colab.research.google.com/drive/1AxlOLc4OiUqdbZig4nqFfqhtZNFdh3ol#scrollTo=cFZgsOb-57bT

Apps190031925 Mail - R...KL-ONLINE-CLASSCoursera | kd-2020-...Global NetAcad Ins...Wiley Certification I...Projects · Dashboar...Online Python Com...

190031925-ds lab 4 .ipynb

FileEditViewInsertRuntimeToolsHelpAll changes saved

Files

sample_data2015 (1).csv2015.csv2016 (1).csv2016.csv2017 (1).csv2017.csv2018 (1).csv2018.csv2019 (1).csv2019.csvgraph.png

```
[14] #countplot
f,ax=plt.subplots(figsize=(10,10))
sns.countplot(data_concat.Region,ax=ax)
plt.xticks(rotation=45)
plt.show()
```

Disk75.75 GB available

Person 1

Goutam

goutamrayavarapugoutam@gmail.com

Sync is off

Turn on sync...

Manage your Google Account

Sign out

Other people

Goutam

Goutam

Guest

+ Add