

# Project Proposal

## Question /needs:

Looking for drugs and where it come from,??

We can extract some information about world drugs production from seized drugs numbers.

There is lots of factors, which can sway numbers on each side such as government drug tolerance act.

But if we look on overall number, we can extract some basic information.

.

## Data description:

Dataset I obtained from <https://www.kaggle.com/ramjasmaurya/drug-seizues-annually-since-1970s/code> in which the data have been obtained from using tool analysis that was used to analysis Drug Seizures annually since 1970s seizures of drugs from 1970s to pre covid period and.

## Tools:

I will be planning to use deep learning model and library of Numby – Pandas – Matloltlip -Seaporn – Python

## MVP Goal:

The goal of this project is to better understand drugs and where it come from, between 2005 to 2019 and which reign is using it and what drug they using:

1- First I just chose what dataset I need from my dataset

2- I import my tools :

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns |
import matplotlib.pyplot as plt
```

3- Import my dataset using :

```
data= pd.read_csv(r"C:\Users\king\Desktop\New folder (25)\Book1.csv")
```

data

```
In [2]: data= pd.read_csv(r"C:\Users\king\Desktop\New folder (25)\Book1.csv")
```

```
In [3]: data
```

```
Out[3]:
```

	Region	Country	ISO Code		Drug Group	Drug	Year	KG Equivalent
0	Asia	Afghanistan	AFG		Opioids	Heroin	2005	7112.400000
1	Asia	Afghanistan	AFG		Opioids	Illicit morphine	2005	1967.000000
2	Asia	Afghanistan	AFG		Opioids	Opium	2005	90990.000000
3	Asia	Afghanistan	AFG		Cannabis-type	Hashish (resin)	2005	42389.000000
4	Asia	Afghanistan	AFG		Opioids	Non-specified opioids	2005	NaN
...	...	...	...	...	...	...	...	...
8042	Asia	Pakistan	PAK	Substances not under international control	Ketamine and phencyclidine-type substances		2019	150.770000
8043	Asia	Pakistan	PAK	Sedatives and Tranquillizers	Benzodiazepines		2019	1.084000
8044	Asia	Pakistan	PAK	Sedatives and Tranquillizers	Other sedatives and tranquillizers		2019	0.985000
8045	Africa	Morocco	MAR		ATS	Ecstasy-type substances	2019	280.318335
8046	Africa	Morocco	MAR		Sedatives and Tranquillizers	Benzodiazepines	2019	4.695000

8047 rows x 7 columns

4- check how many data column I have:

```
In [4]: data.columns
```

```
Out[4]: Index(['Region', 'Country', 'ISO Code', 'Drug Group', 'Drug', 'Year',
              'KG Equivalent'],
              dtype='object')
```

5- In this code I check which Region we are using:

```
data.Region.value_counts ()
```

```
In [5]: data.Region.value_counts()

Out[5]: Asia      5962
Africa    2085
Name: Region, dtype: int64
```

6- In this I checked which value I was using in every country:  
data.Country.value\_counts(50)

```
In [6]: data.Country.value_counts(50)

Out[6]: China, Hong Kong SAR      0.040015
Japan      0.033677
India      0.032310
Armenia     0.031440
China, Taiwan Province of China  0.027588
...
Maldives    0.000746
Lesotho     0.000497
Seychelles  0.000373
Rwanda      0.000249
Democratic Republic of the Congo  0.000124
Name: Country, Length: 93, dtype: float64
```

7- I use this to groupby Drug Group ,Region, KG Equivalent:

```
df=pd.read_csv(r"C:\Users\king\Desktop\New folder
(25)\Book1.csv")
new = df.groupby(['Drug Group','Region']).agg({'KG Equivalent':
'sum'})
# Change: groupby state_office and divide by sum
prct = new.groupby(level=0).apply(lambda x:
round(100 * x /
float(x.sum()),2)).reset_index()
```

```
In [7]: df=pd.read_csv(r"C:\Users\king\Desktop\New folder (25)\Book1.csv")
new = df.groupby(['Drug Group','Region']).agg({'KG Equivalent': 'sum'})
# Change: groupby state_office and divide by sum
prct = new.groupby(level=0).apply(lambda x:
round(100 * x / float(x.sum()),2)).reset_index()
```

8- I use this next command to see my head :  
prct.head(111)

```
In [8]: prct.head(111)
```

```
Out[8]:
```

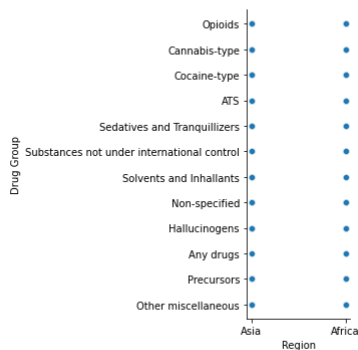
	Drug Group	Region	KG Equivalent
0	ATS	Africa	6.45
1	ATS	Asia	93.55
2	Any drugs	Africa	NaN
3	Any drugs	Asia	NaN
4	Cannabis-type	Africa	42.98
5	Cannabis-type	Asia	57.02
6	Cocaine-type	Africa	54.27
7	Cocaine-type	Asia	45.73
8	Hallucinogens	Africa	2.20
9	Hallucinogens	Asia	97.80
10	Non-specified	Africa	38.25

9- This function provides access to several different axes-level functions that show the relationship between two variables with semantic mappings of subsets in this :

```
sns.relplot(x="Region",y="Drug Group",data=data)
```

```
In [9]: sns.relplot(x="Region",y="Drug Group",data=data)
```

```
Out[9]: <seaborn.axisgrid.FacetGrid at 0x1636f1d9a20>
```

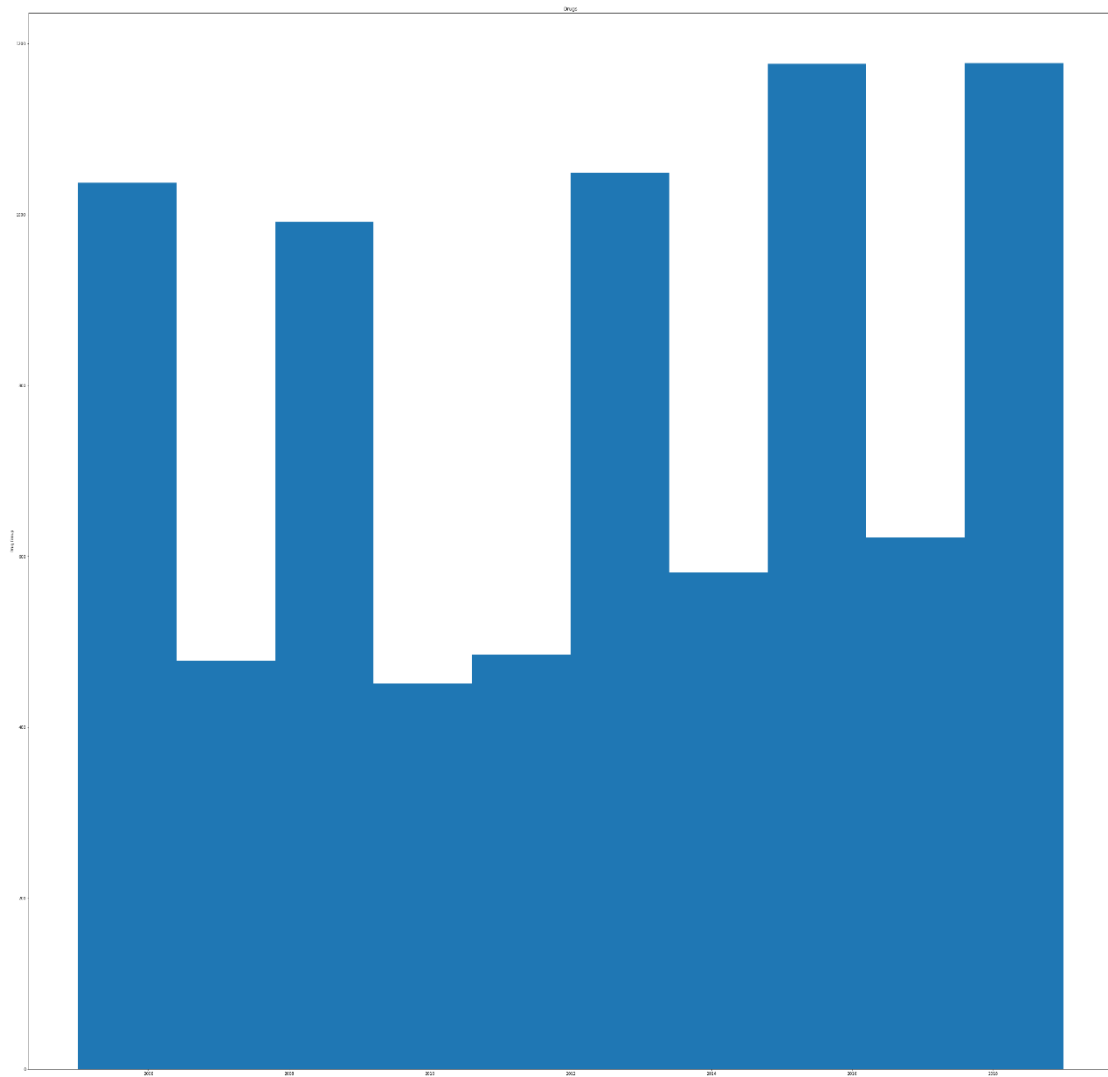


```
7- fig, ax = plt.subplots(2, 1)
```

9- I use figsize to see the drug by year and by drug group :

```
plt.figure(figsize=(50,50))
```

```
plt.hist(data['Year'])  
plt.title("Drugs")  
plt.xlabel("Year")  
plt.ylabel("Drug Group")  
plt.show()
```



12- Show the counts of observations in each categorical bin using bars

A count plot can be thought of as a histogram across a categorical, instead of quantitative, variable. The basic API and options are identical to those for `barplot()`, so you can compare counts across nested variables:

```
sns.set_style('whitegrid')
plt.figure(figsize=(35,15))
sns.countplot(x="Drug Group", data=data, palette='cubehelix');
```

---

```
sns.set_style('whitegrid')
plt.figure(figsize=(35,15))
sns.countplot(x="Year", data=data, palette='icefire')
```

