A Look at Drug Use and Overdose Data

Cem Tener, Luis Weekes

Aim

- Ask questions.
- Provide insight.
- Inform public policy makers.

Drug Use by Age

Substance_Use = Percentage of those in an age group who used said substance in the past year.

Substance_Freq = Median number of times users in an age group used said substance in the past year.

Data From: “National Survey on Drug Use and Health from the Substance Abuse and Mental Health Data Archive”

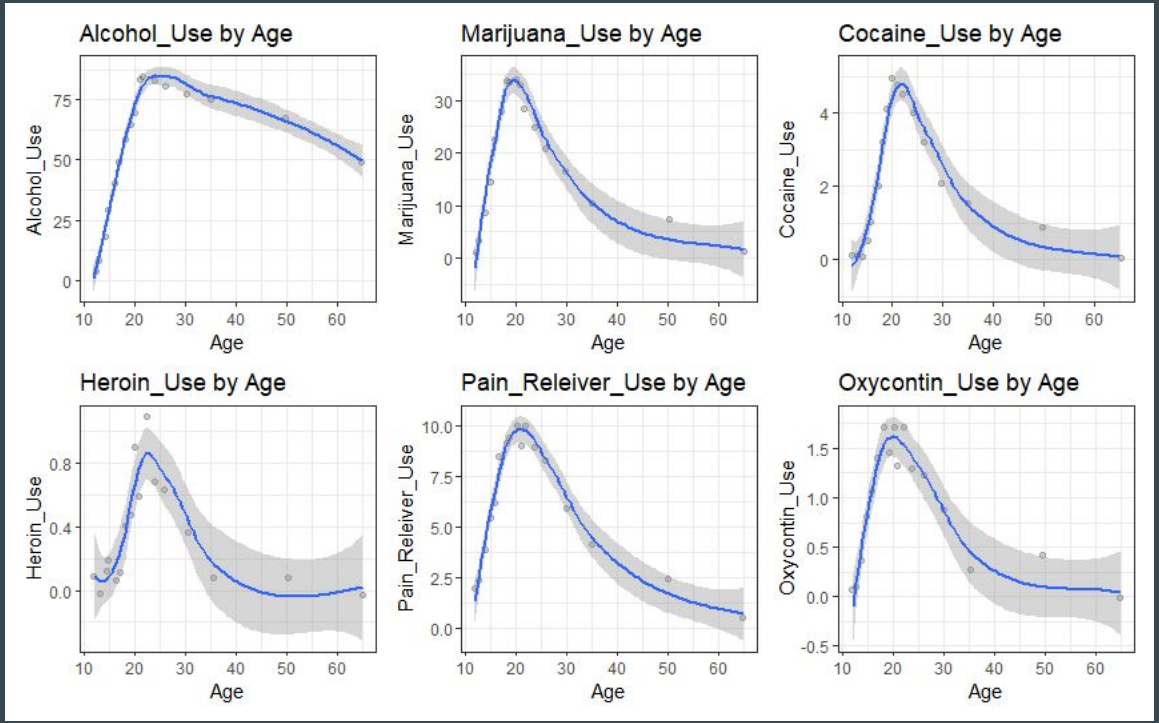
Cleaning and Wrangling

```
drug_use_by_age_data <- read.csv("~/R programming class/drug_use_by_age_data.csv")

Drug_use_by_age<-drug_use_by_age_data%>%
  mutate(
    Alcohol_Use=alcohol.use,
    Alcohol_Freq=alcohol.frequency,

    Marijuana_Use=marijuana.use,
    Marijuana_Freq=marijuana.frequency,

    Cocaine_Use=cocaine.use,
    Cocaine_Freq=as.numeric(cocaine.frequency),

    Heroin_Use=heroin.use,
    Heroin_Freq=as.numeric(heroin.frequency),

    Pain_Releiver_Use=pain.releiver.use,
    Pain_Releiver_Freq=pain.releiver.frequency,

    Oxycontin_Use=oxycontin.use,
    Oxycontin_Freq=as.numeric(oxycontin.frequency),

    Age=age

  )
```

How many people in each age group have taken the substance?

```
Average_use_by_age_plot<-function(.data, y_var){
  ggplot(.data, aes(x = Age)) +
    aes_string(y = y_var)+
    geom_jitter(alpha = 0.2)+
    geom_smooth() +
    labs(
      title = paste( y_var,"by Age ")
    )+
    theme_bw()
   }


c("Alcohol_Use", "Marijuana_Use", "Cocaine_Use", "Heroin_Use", "Pain_Releiver_Use","Oxycontin_Use")%>%
  map(Average_use_by_age_plot, .data = Drug_use_by_age)%>%
  patchwork::wrap_plots(ncol = 3)
```

Plots of Percent Using per Age Group



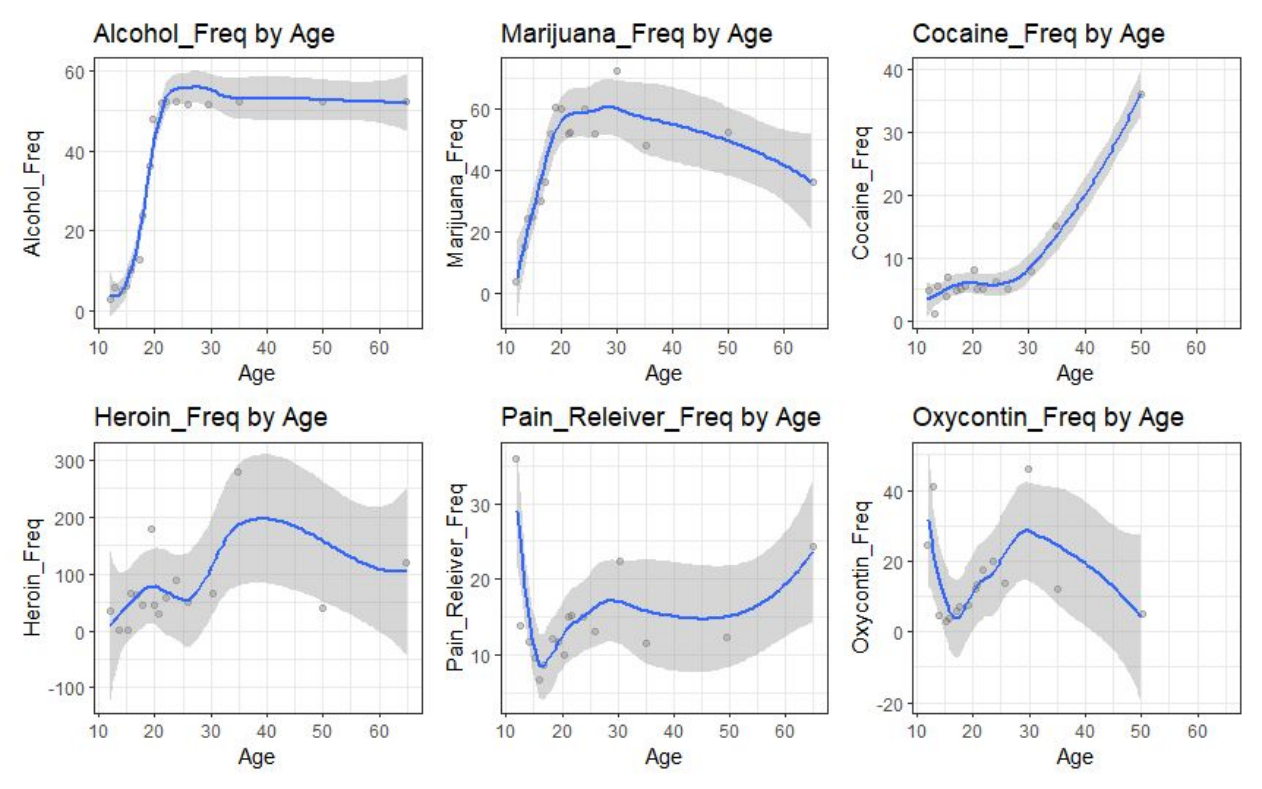
What does this mean?

- Policy makers might now have a better idea of which age groups have tried a particular drug the most.
- What prompts people from that age group to try that drug?

Which age group uses the substance the most?

```
Freq_of__use_by_age_plot<-function(.data, y_var){
  ggplot(.data, aes(x = age)) +
    aes_string(y = y_var)+
    geom_jitter(alpha = 0.2)+
    geom_smooth() +
    scale_x_continuous(name = "Age", limits=c(10, 65))+
    labs(
      title = paste( y_var,"by Age "),
      #caption = "National Survey on Drug Use and Health from the Substance Abuse and Mental Health Data Archive"
  )
}

c("Alcohol_Freq", "Marijuana_Freq", "Cocaine_Freq", "Heroin_Freq", "Pain_Releiver_Freq","Oxycontin_Freq")%>%
  map(Average_use_by_age_plot, .data = Drug_use_by_age)%>%
  patchwork::wrap_plots(ncol = 3)
```

Plot of Frequency of Use by Age



What does this mean?

- Frequency stays the same or increases for the most part.
- No data points past 50 for Cocaine and Oxycontin.
- Which age group is most addicted?

Drug Overdoses by State

Data from: “CDC Injury Center Drug Overdose Deaths”

- Interesting cleaning issue.

	V1	V2	V3	V4	V
1	State	State Abbreviation	2019 Age-adjusted Rate (per 100,000 population)	2019 Number of Deaths	2
2	Alabama	AL	16.3	768	1
3	Alaska	AK	17.8	132	1
4	Arizona	AZ	26.8	1907	2
5	Arkansas	AR	13.5	388	1
6	California	CA	15	6198	1

```r
Drug_deaths_by_state<-Drug_deaths_by_state%>%
  select(V1, V2, V3, V17, V21)%>%
  filter(V1!="State")%>%
  mutate(
  State= V1,
  State_abbr=V2,
  Age_adjusted_death_rate_2019=V3,
  Poverty_rate= V17,
  Urban_pop_percentage= V21
  )%>%
  select( State,  State_abbr, Age_adjusted_death_rate_2019, Poverty_rate, Urban_pop_percentage )%>%
  mutate(Age_adjusted_death_rate_2019 = as.numeric(Age_adjusted_death_rate_2019),
         Poverty_rate = as.numeric(Poverty_rate),
         Urban_pop_percentage = as.numeric(Urban_pop_percentage)
         )
```

What do you think?

Will states with higher poverty rates have more drug related deaths?

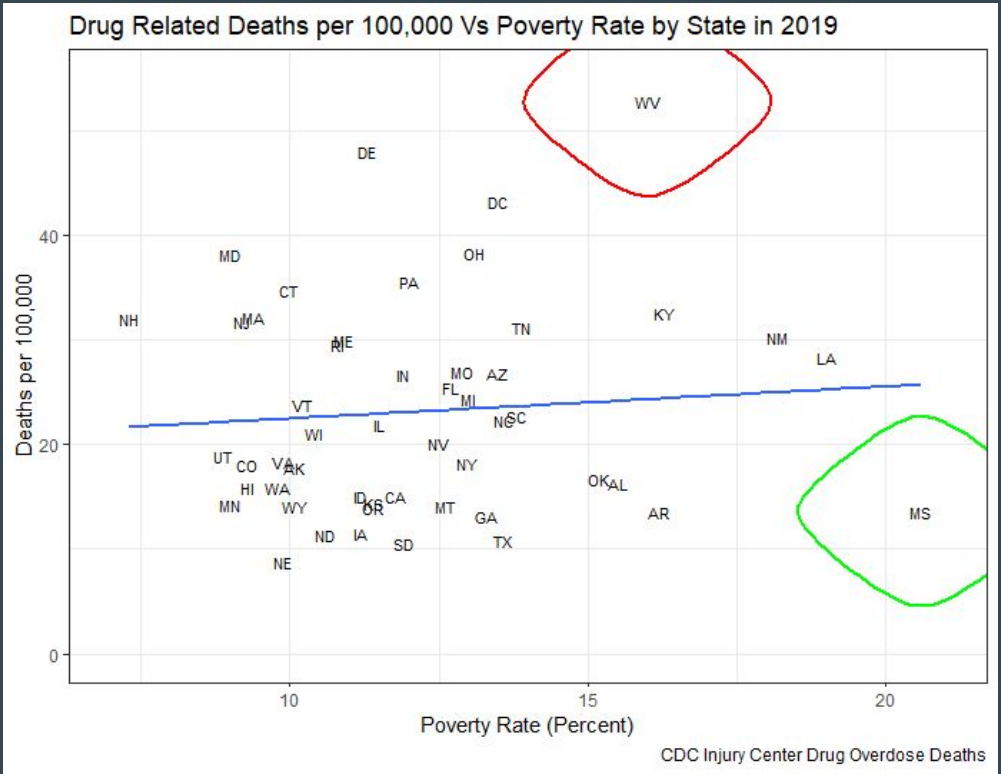
and

Will states with a larger urban population have more drug related deaths?

Poverty Rate and Drug Related Deaths by State

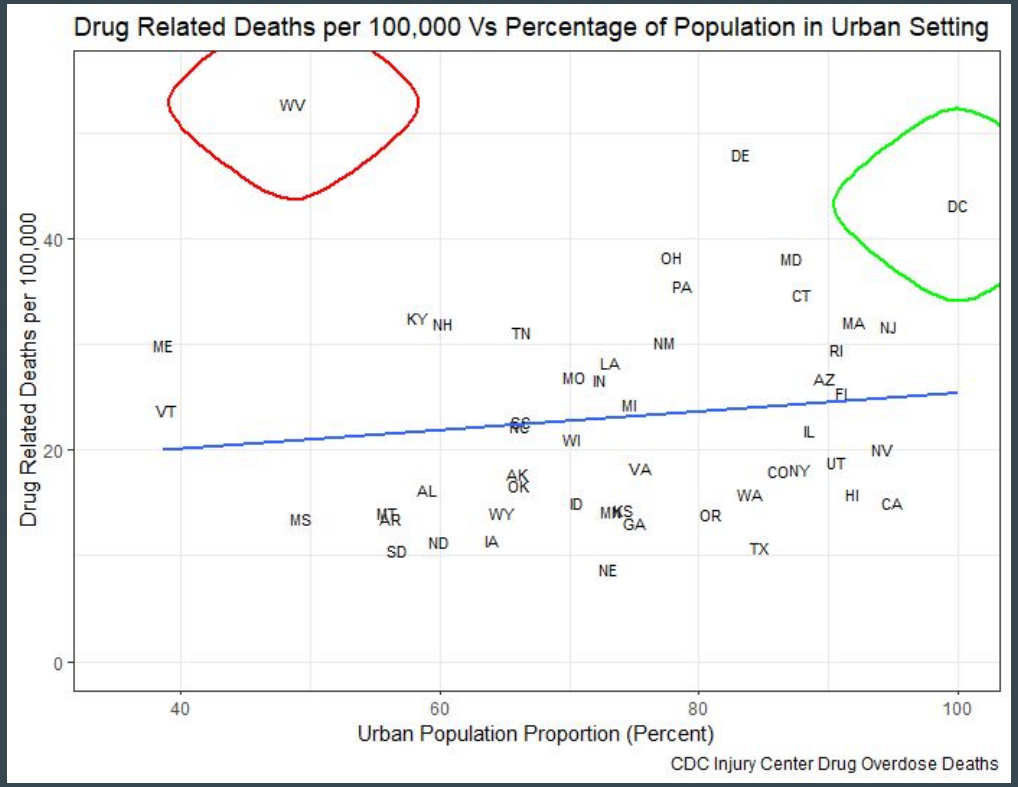
```
West_Virginia<-Drug_deaths_by_state%>%
  filter(State_abbr== "WV")

Mississippi<-Drug_deaths_by_state%>%
  filter(State_abbr== "MS")

D_C<-Drug_deaths_by_state%>%
  filter(State_abbr== "DC")

library(ggalt)

Death_rate_vs_Poverty_rate_plot<-ggplot(data= Drug_deaths_by_state, aes(y = Age_adjusted_death_rate_2019, x = Poverty_rate))+
  geom_text(aes(label = State_abbr),size=3)+
  scale_x_continuous(name = "Poverty Rate (Percent)", limits=c(7, 21))+
  scale_y_continuous(name = "Deaths per 100,000", limits=c(0, 55))+
  geom_smooth(method="lm", se= FALSE)+
  labs(
    title = "Drug Related Deaths per 100,000 Vs Poverty Rate by State in 2019",
    caption = "CDC Injury Center Drug Overdose Deaths"
  )+
  theme_bw()+
  geom_encircle(aes(x=Poverty_rate, y=Age_adjusted_death_rate_2019),
                data=West_Virginia,
                color="red",
                size=2)+
  geom_encircle(aes(x=Poverty_rate, y=Age_adjusted_death_rate_2019),
                data=Mississippi,
                color="green",
                size=2)
```

Plot of Deaths per 100,000 Against Poverty Rate




```
data:  x and y
t = 0.57786, df = 49, p-value = 0.566
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.1977969  0.3499214
sample estimates:
       cor
0.08227132
```

Plot of Deaths per 100,000 Against Urban Population Percentage

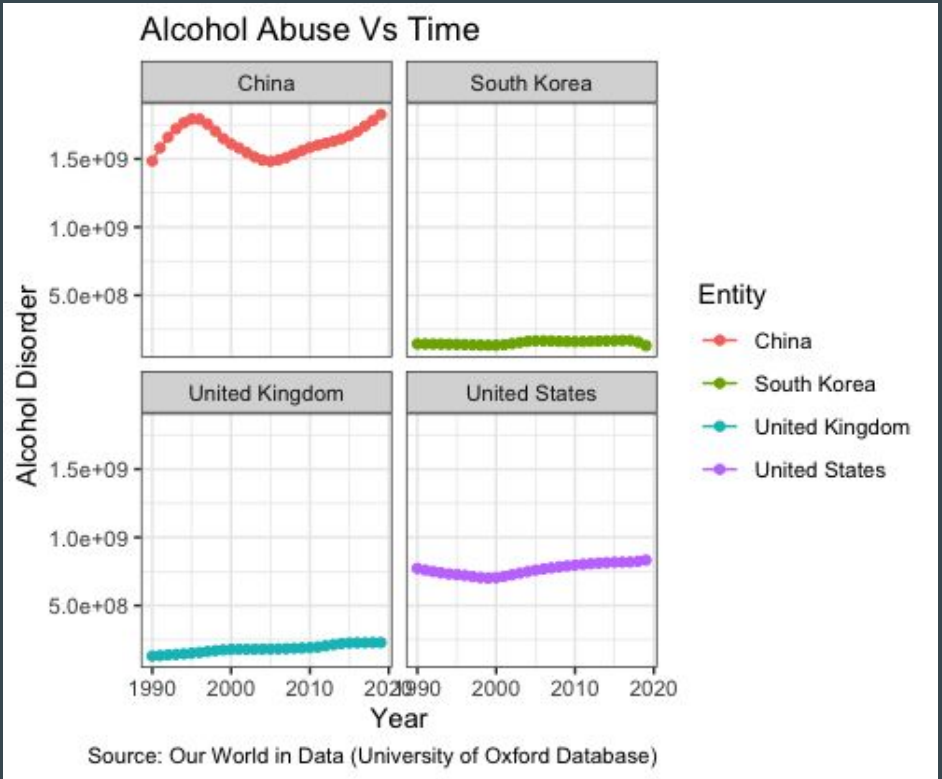



```
data:  x and y
t = 0.91095, df = 49, p-value = 0.3668
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.1519395  0.3907354
sample estimates:
      cor
0.1290476
```

What does this mean?

- Drug related problems occur in very different places.
- Policy makers must not ignore certain types of states based on preconceived notions.

How did alcohol disorders evolve over time?



Can it be due to economic growth?

- China saw continuous real GDP growth of at least 5% during 90’s.
- In 1996, Chinese economy continued to grow at about 9.5% with low inflation.
- Do economics growth affect alcohol disorders?

Getting GDP and Disorder Data Together

```
alcohol_drug_country<-alcohol_country %>%
  rename(alcohol_use_disorder_percent =
"Prevalence...Alcohol.use.disorders...Sex..Both...Age..Age.standardized..Percent.",
         drug_use_disorder_percent="Prevalence...Drug.use.disorders...Sex..Both...Age..Age.stan
dardized..Percent.",
         population="Population..historical.estimates."
         ) %>%
  mutate(alcohol_disorder = (population*alcohol_use_disorder_percent),
         drug_disorder = (population*drug_use_disorder_percent))%>%
  filter(Entity == "United States" |Entity ==  "South Korea" |Entity ==  "China" |Entity ==
"United Kingdom") %>%
    select(Entity, Year,alcohol_disorder,drug_disorder)%>%
  filter(Year>1970)
alcohol_drug_country<-na.omit(alcohol_drug_country)
```

```
#Gdp per capita data for each country
us_gdp_capita=read.csv(file=file.choose())
us_gdp_capita<-us_gdp_capita%>%
  mutate(Year=lubridate::year(DATE),
         Entity = "United States",
         gdp_per_capita=A939RX0Q048SBEA)%>%
  select(Year,Entity,gdp_per_capita)
us_gdp_capita
```

```
#Combine all the datasets
full_gdp_data<-rbind(korea_gdp_capita,china_gdp_capita,us_gdp_capita,uk_gdp_capita)
full_gdp_data

full_data<-full_join(full_gdp_data,alcohol_drug_country,c("Year"="Year","Entity"="Entity"))
full_data
```

```
> head(full_data)
  Year      Entity gdp_per_capita alcohol_disorder drug_disorder
1 1990 South Korea       9365.395        143756031      35567432
2 1991 South Korea      10272.621        143575360      36000735
3 1992 South Korea      10796.605        142768035      36307723
4 1993 South Korea      11422.515        141909698      36631566
5 1994 South Korea      12356.278        141038787      36910428
6 1995 South Korea      13408.681        140146984      37217158
```

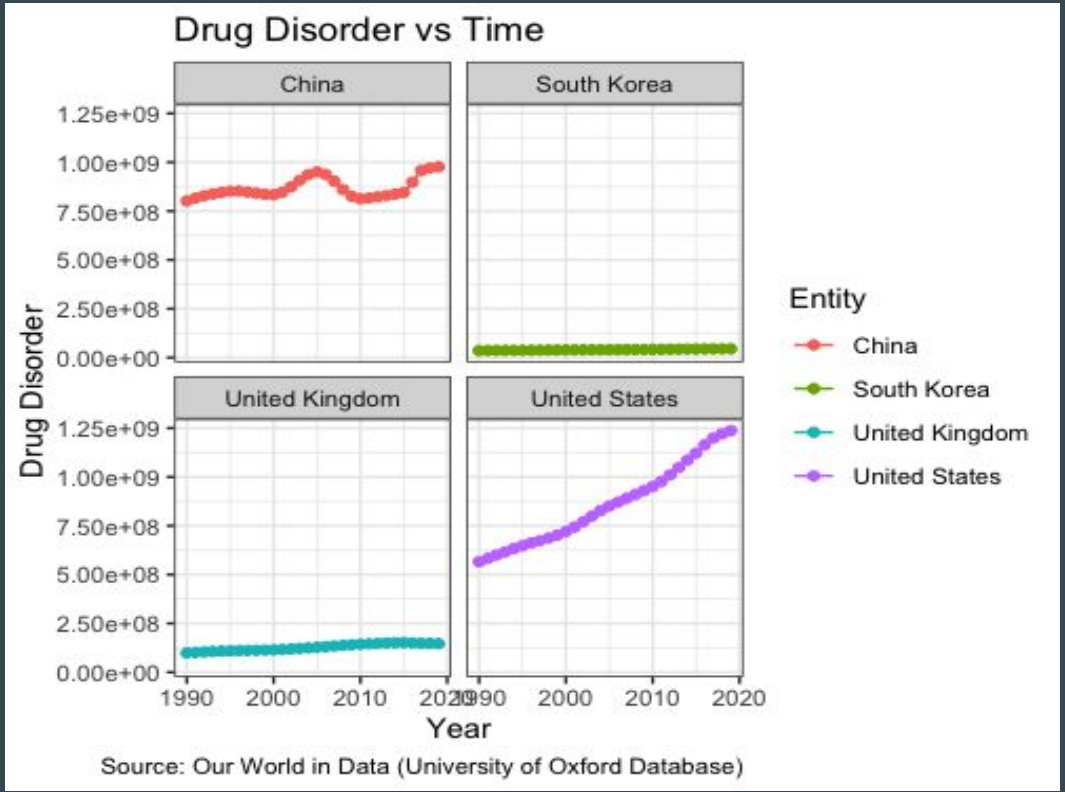
No correlation between GDP per capita and alcohol disorder

Correlation test for alcohol disorders and GDP per capita
t= 1.292, df = 27, p-value = 0.2073

95 percent confidence interval:
(-0.1373483 - 0.5584234)

sample estimates:
0.2413044

How did the drug disorders evolve over time?



Does economic growth affect drug disorders in the US and China?

Correlation test for drug disorders and GDP per capita for China

t = -0.059496, df = 27, p-value = 0.953

95 percent confidence interval:
(-0.3763755 0.3565526)

Correlation estimate: -0.01144919

Correlation test for drug disorders and GDP per capita for the US

t = -0.27952, df = 27, p-value = 0.782

95 percent confidence interval:
(-0.4121089 - 0.3190708)

Correlation estimate: -0.05371661

We cannot conclude that economic growth is an important factor affecting alcohol and drug disorders.

What does this mean?

- Policy makers should help people who are struggling with alcohol and drug disorders regardless of how the economy is doing.
- It disproves the idea that people might only need help in varying economic times.