

EXPLORATIVE DATA ANALYSIS

THIS PROJECT IS CONCENTRATED ON EXPLORATIVE DATA ANALYSIS,
EDA. EDA IS AN ESSENTIAL PART OF THE DATA ANALYSIS PIPELINE.

OBJECTIVE

- Investigate drug abuse among different age groups.
- To understand the relationship between variables if any.
- Determine which age group to pay special attention to for drug abuse.

DATA SCIENCE/ANALYSIS PIPELINE

- Obtaining the data: This is the first approach
- Scrubbing or cleaning the data: This includes checking for missing data and converting data to its right format.
- Exploring the data: This involves understanding the variables and finding patterns in the dataset. This is what this project is about.
- Modeling the data: This is where the predictive power on whether or not something happens is exercised.
- Interpreting the data: This is the part where conclusions are made.

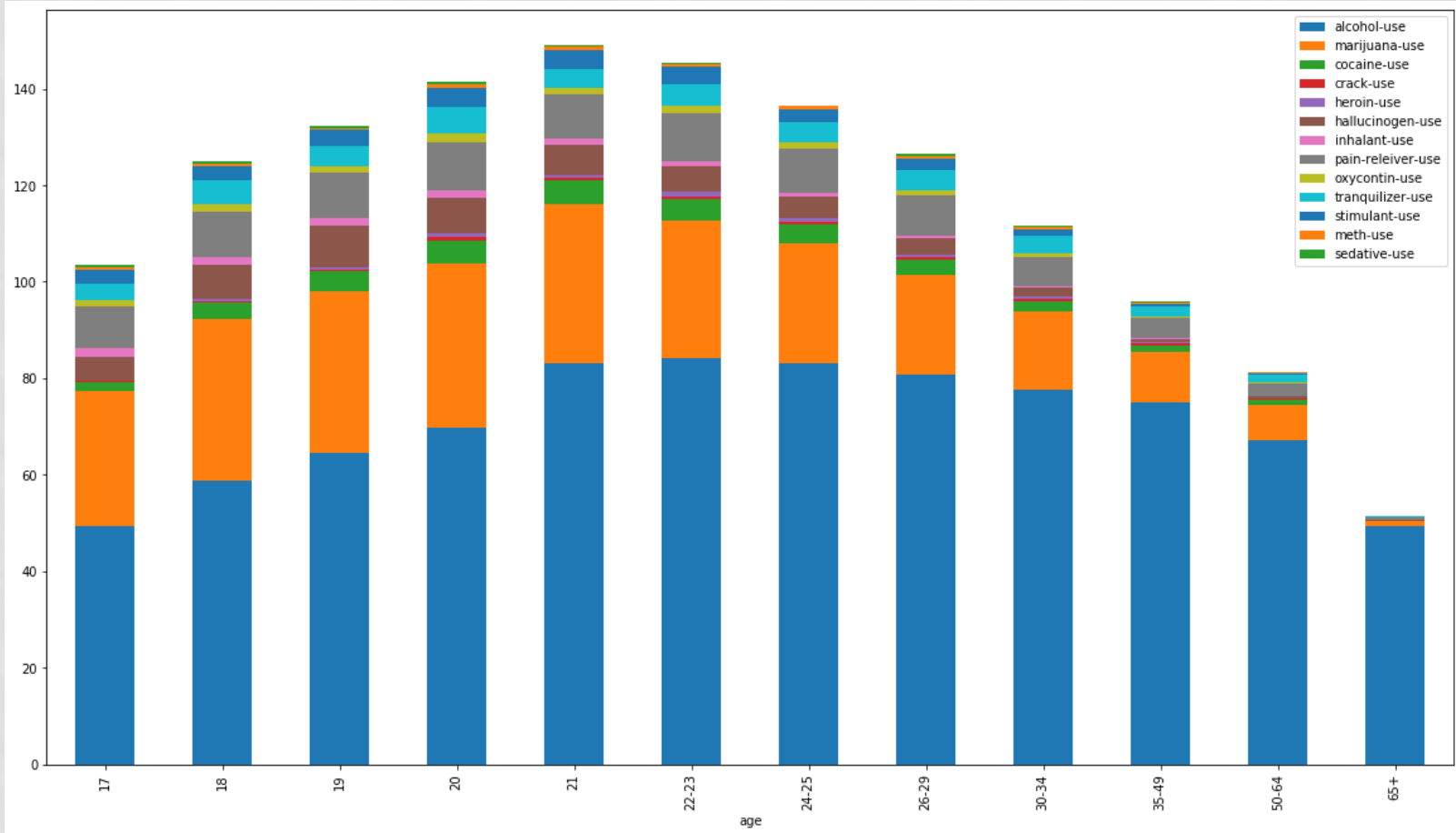
PROBLEM

- One of the most common problems among all age group is drug abuse. This spans across all kinds of drugs be it pain reliever, alcohol, sedatives etc.
- The rate at which youths nowadays abuse their intake of alcohol and the likes is not exactly unprecedented but it is quite alarming. I think because most of them do not have good role models.

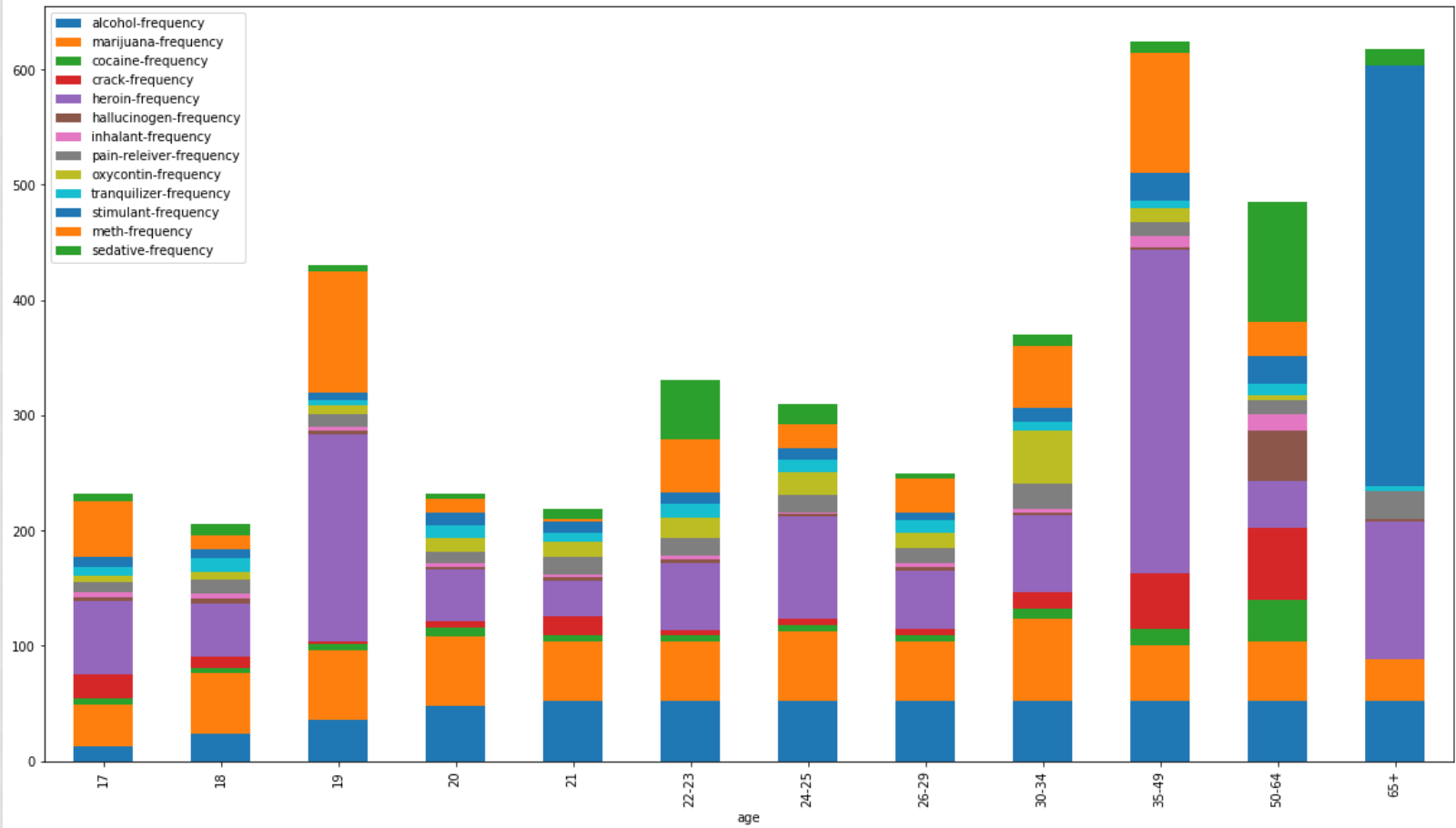
THE DATASET

- 'age','alcohol-use','marijuana-use','cocaine-use','crack-use','heroin-use','hallucinogen-use','inhalant-use','pain-releiver-use','oxycontin-use','tranquilizer-use','stimulant-use','meth-use','sedative-use','alcohol-frequency','marijuana-frequency','cocaine-frequency','crack-frequency','heroin-frequency','hallucinogen-frequency','inhalant-frequency','pain-releiver-frequency','oxycontin-frequency','tranquilizer-frequency','stimulant-frequency','meth-frequency','sedative-frequency'. These are the drugs and the frequency at which they are being used in all ages.

DRUG ABUSE AMONG AGE GROUPS



DRUG FREQUENCY AMONG AGE GROUP



EXPLANATION

- From the above charts, it appears that the abuse of drugs peaked around the age groups between 18-23 years old. This is around the time when parents tend to give their kids more freedom.
- Pain relievers is among the highest drugs being abused; followed by hallucinogen
- Also, the charts implied that pain reliever is the most abused drug among all age groups after alcohol and marijuana.

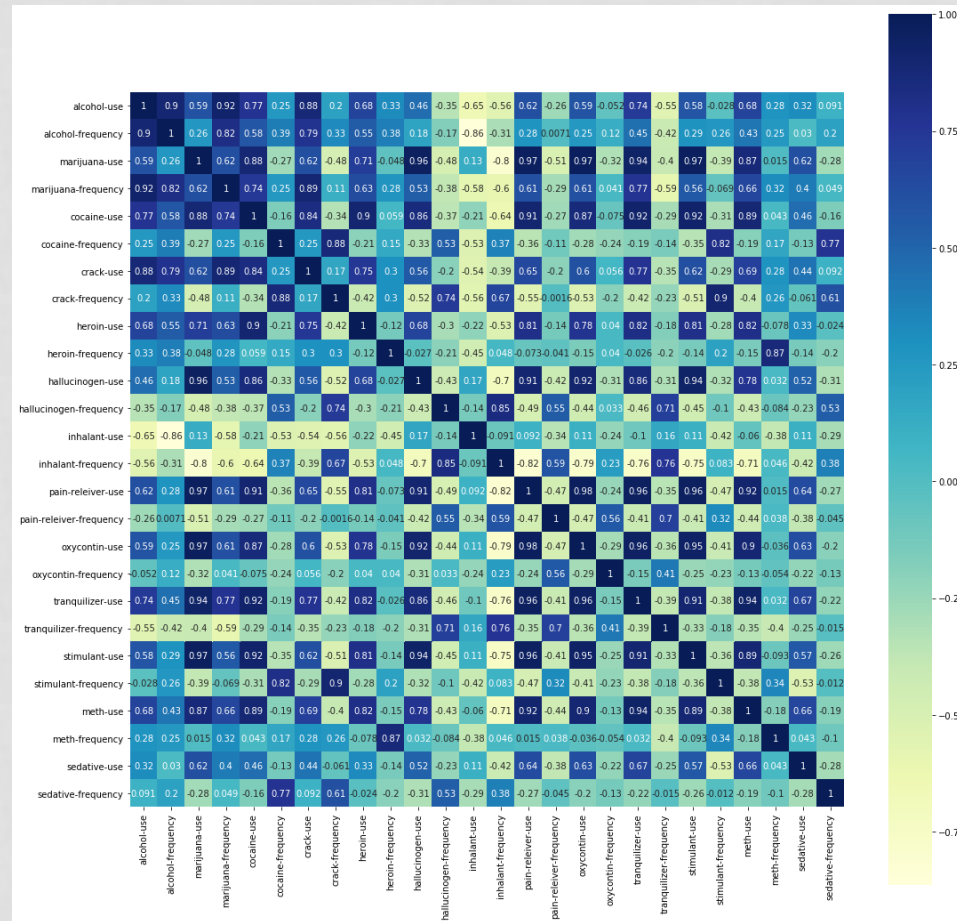
EXPLANATION

- From the drug frequency chart, heroin users tend to abuse it repeatedly.
- Age group to monitor would be “19”, the age group “35-39” and “65+”.

CORRELATION RELATIONSHIP

- To see if there is correlation between various drugs, heatmap with correlation coefficients is used.
- See next page for heatmap

CORRELATION TABLE FOR DRUGS AND THEIR FREQUENCY

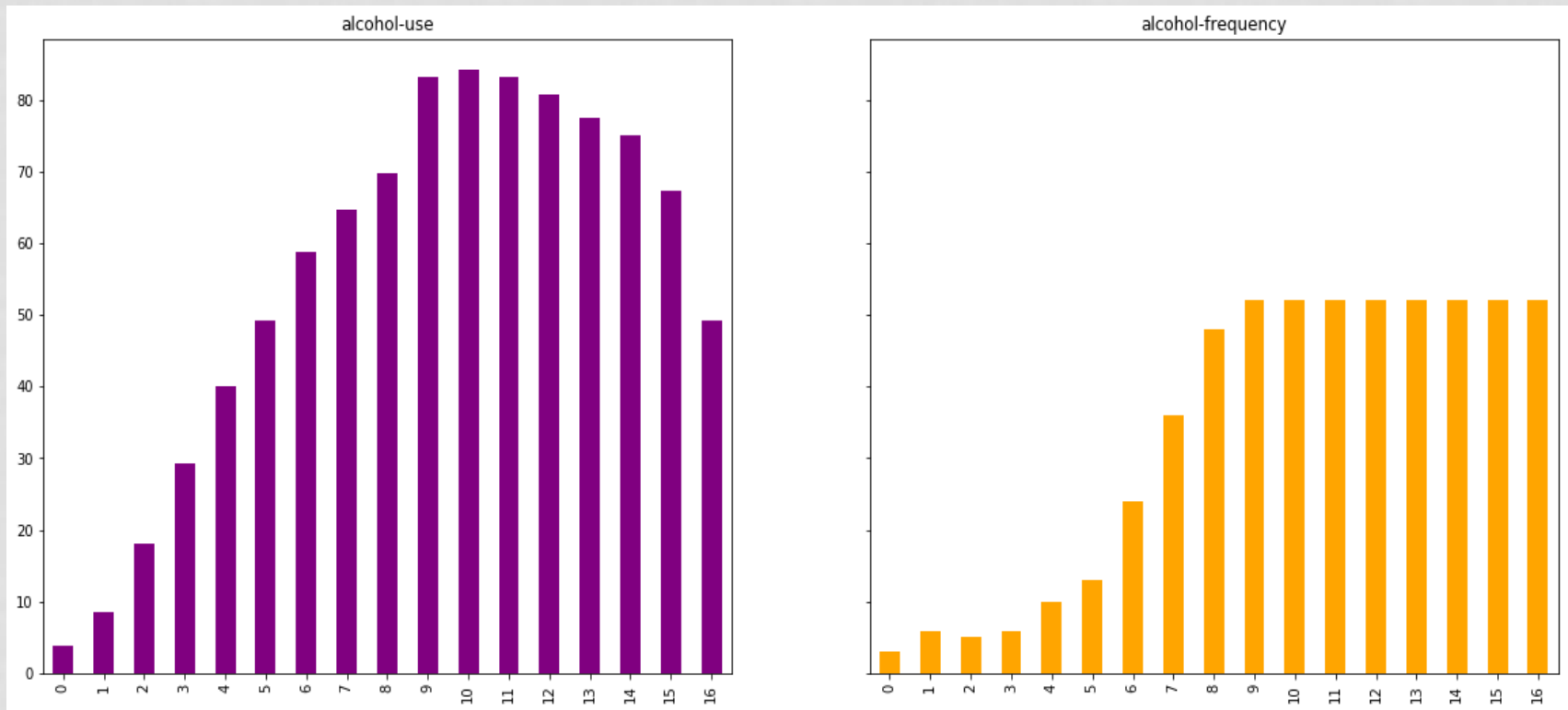


CORRELATION RELATIONSHIP

- There are positive correlation between some drugs usage and their frequency. Take for example, alcohol use and marijuana use have a correlation coefficient of 0.59 and alcohol frequency and marijuana frequency have a correlation coefficient of 0.82.
- Any correlation coefficient lower than 0.50 is a negative correlation and 0.50 and above for positive correlation.

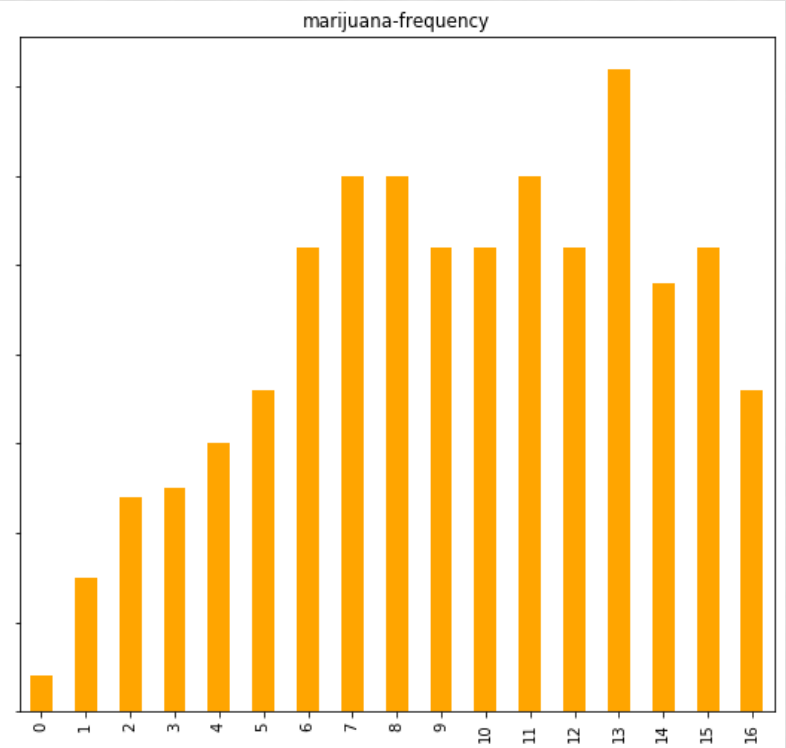
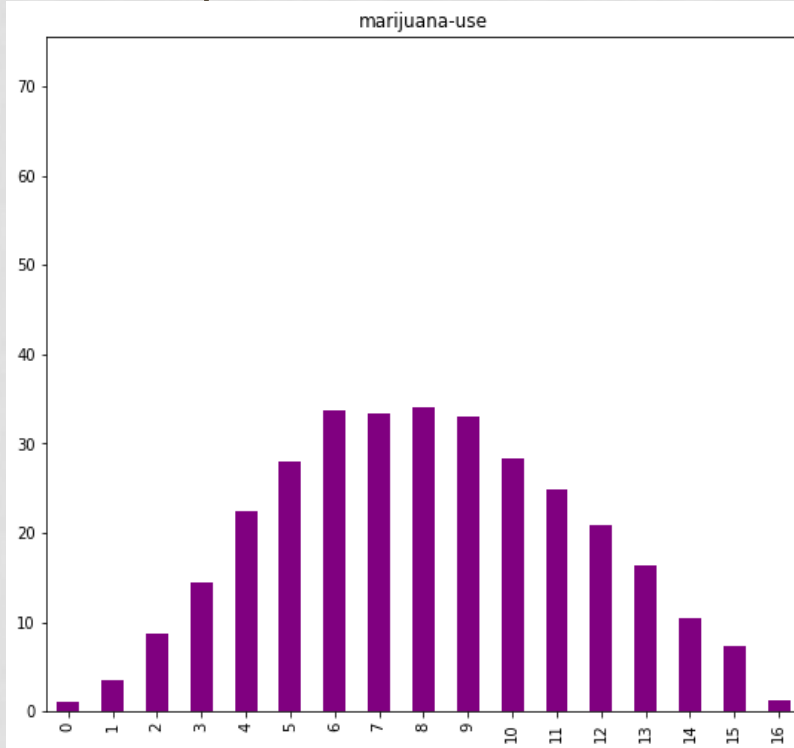
DRUGS AND THEIR FREQUENCY

- Alcohol



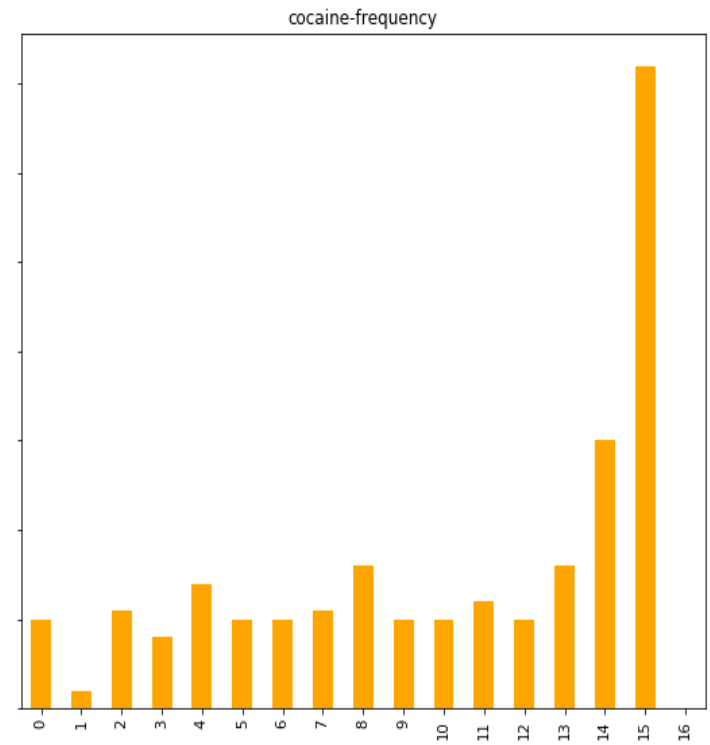
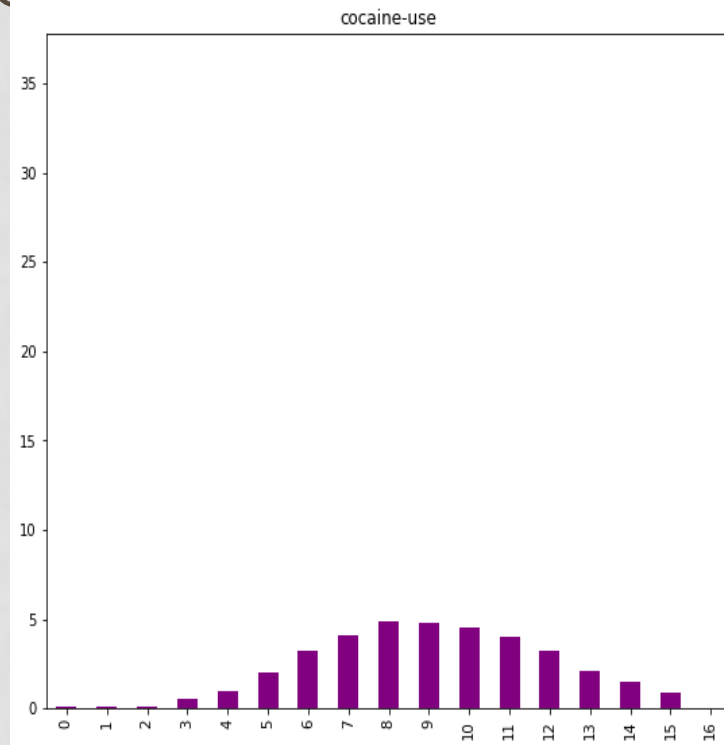
DRUGS AND THEIR FREQUENCY

- Marijuana



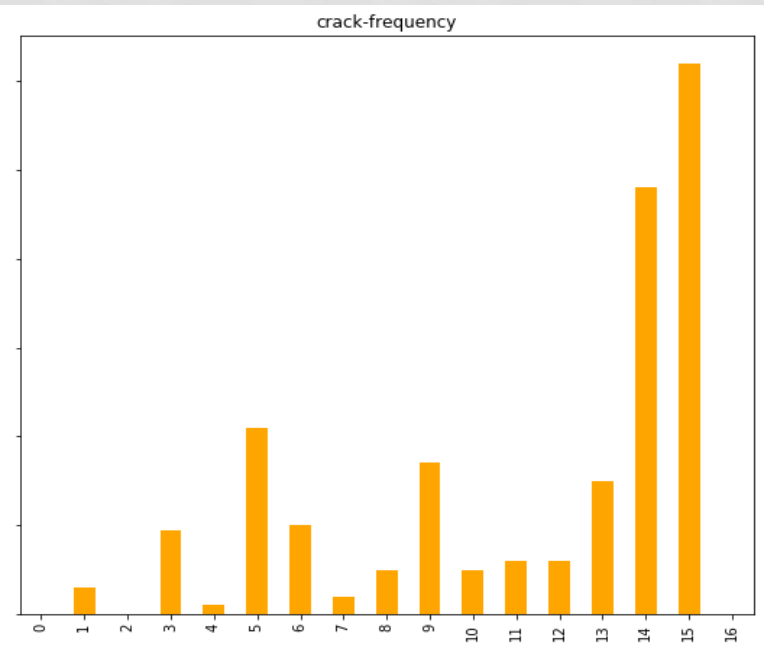
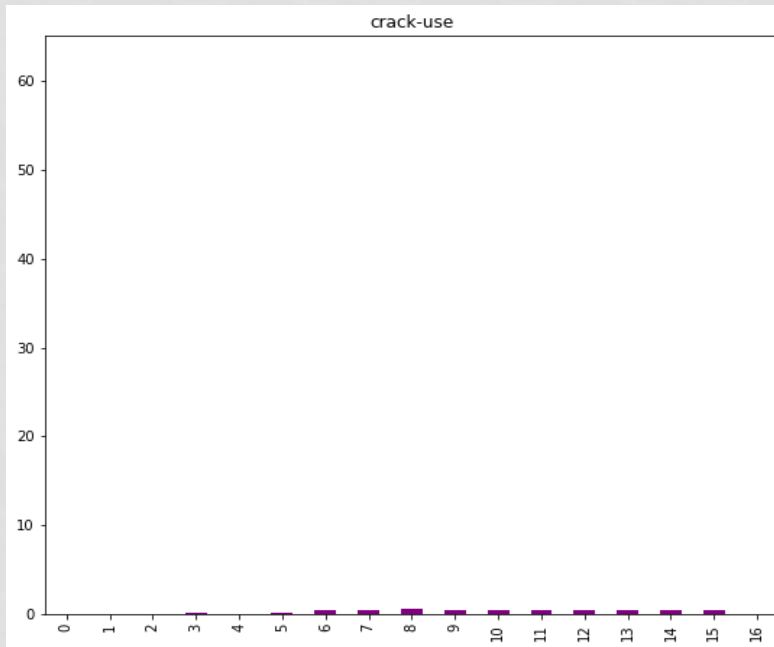
DRUGS AND THEIR FREQUENCY

- Cocaine



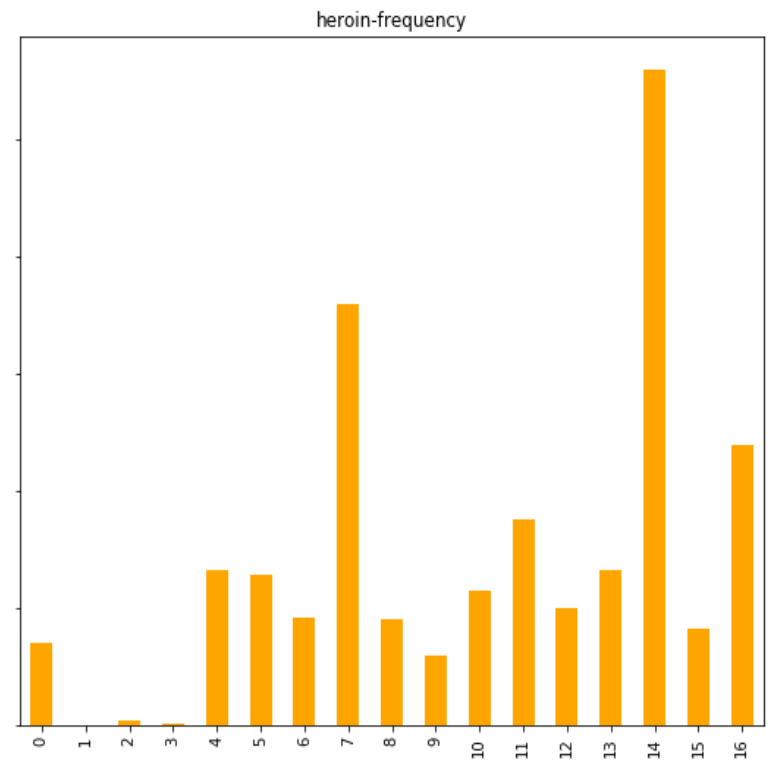
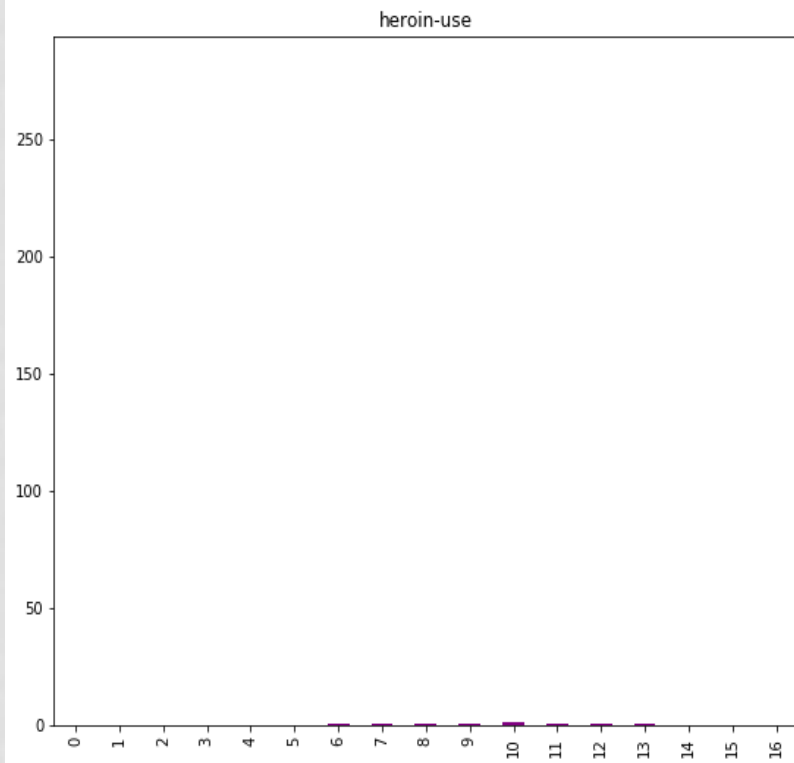
DRUGS AND THEIR FREQUENCY

- Crack



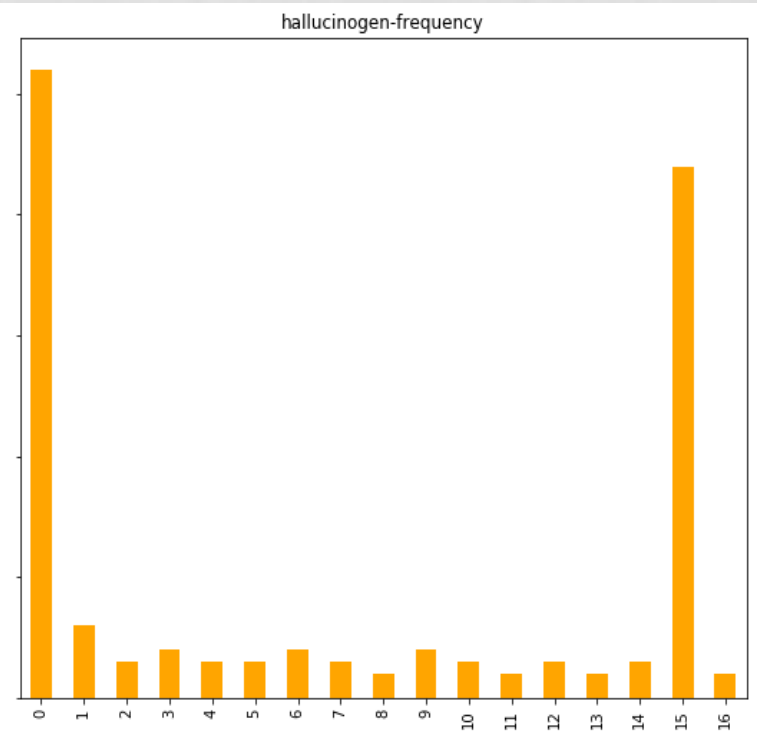
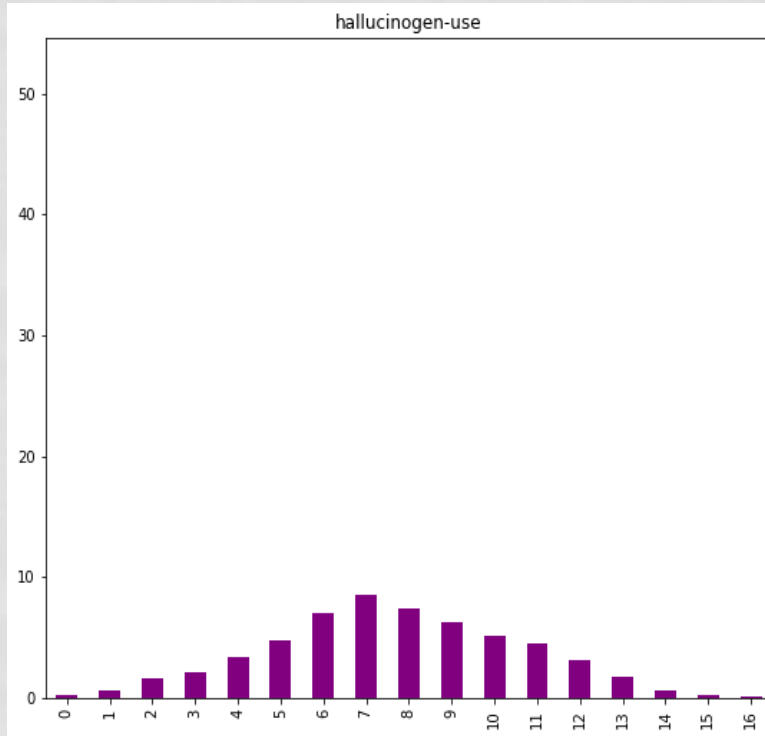
DRUGS AND THEIR FREQUENCY

- Heroin



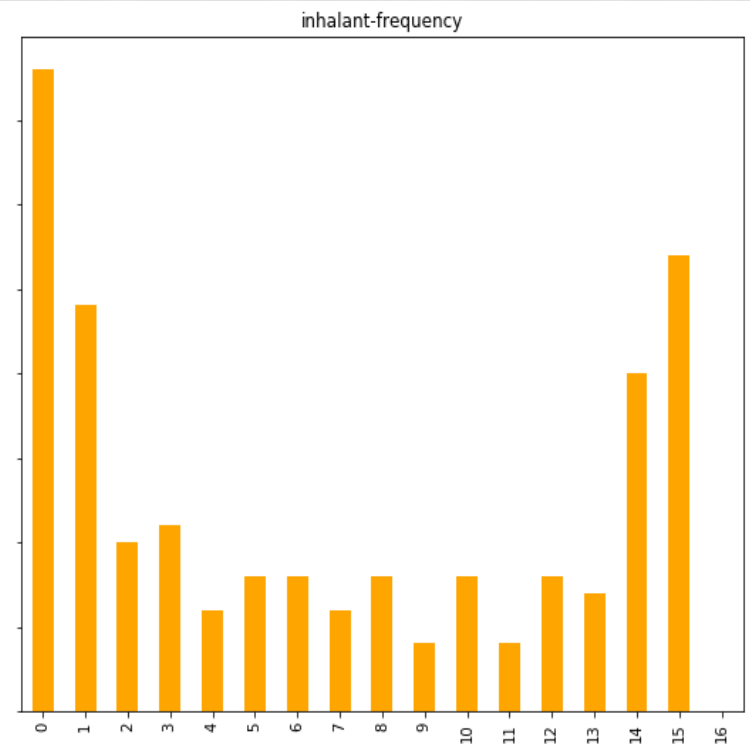
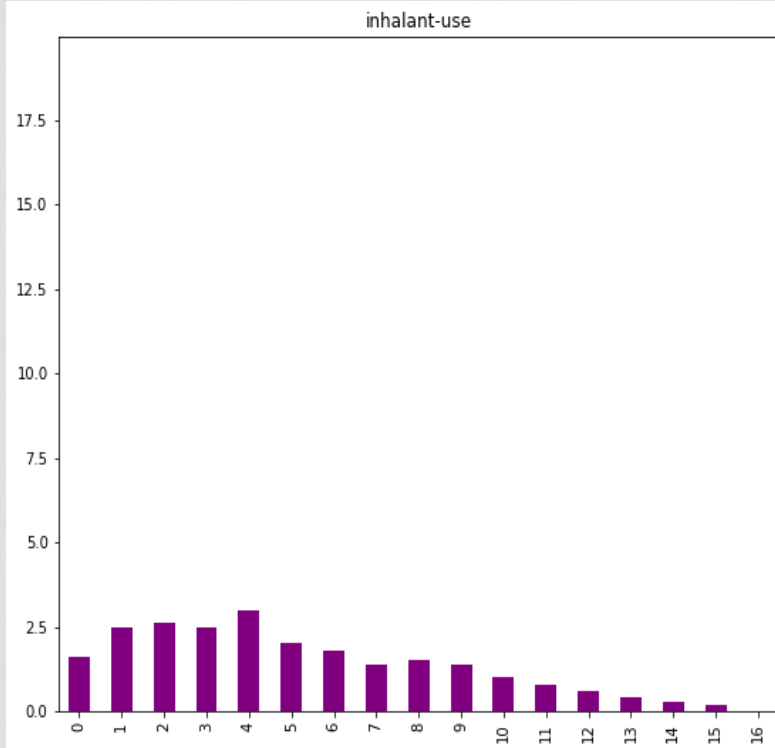
DRUGS AND THEIR FREQUENCY

- Hallucinogen



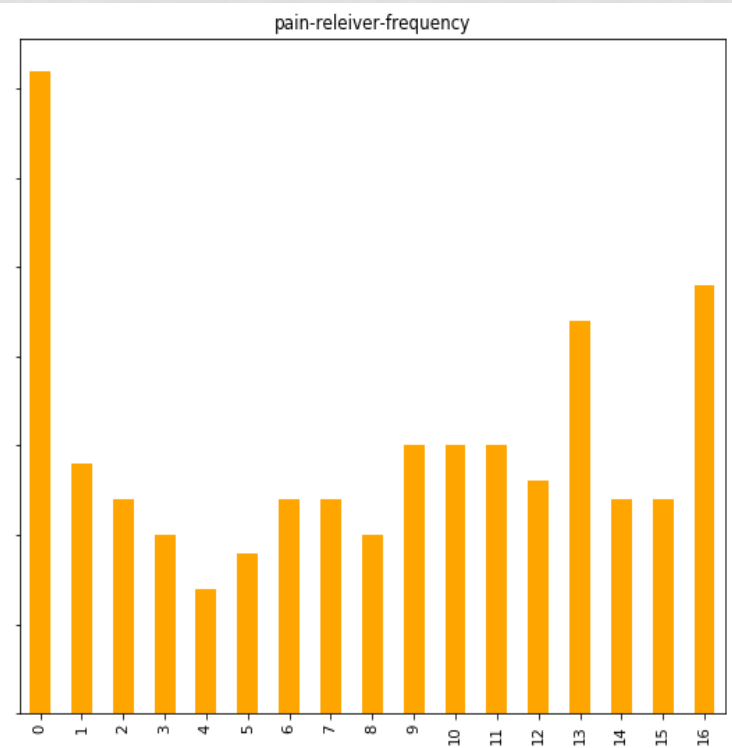
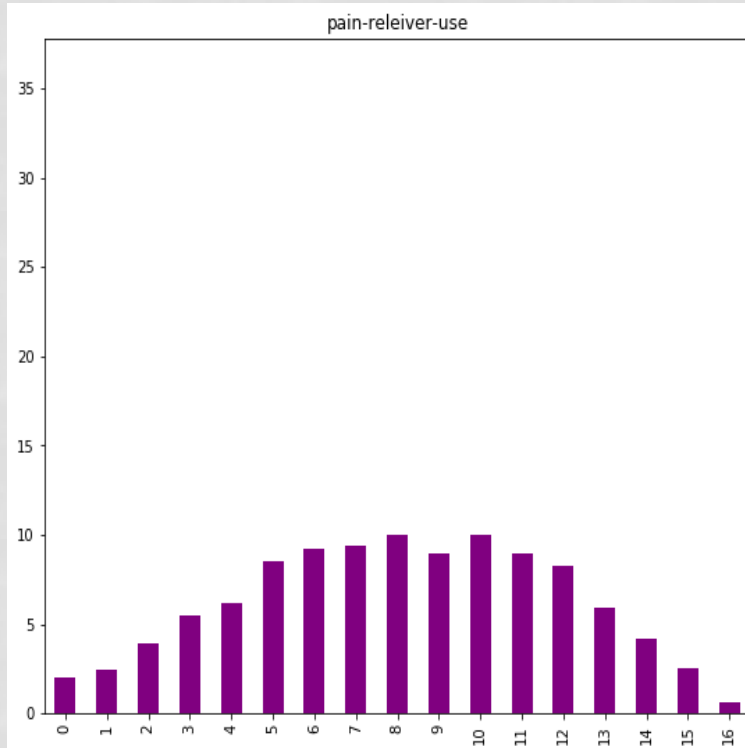
DRUGS AND THEIR FREQUENCY

- Inhalant



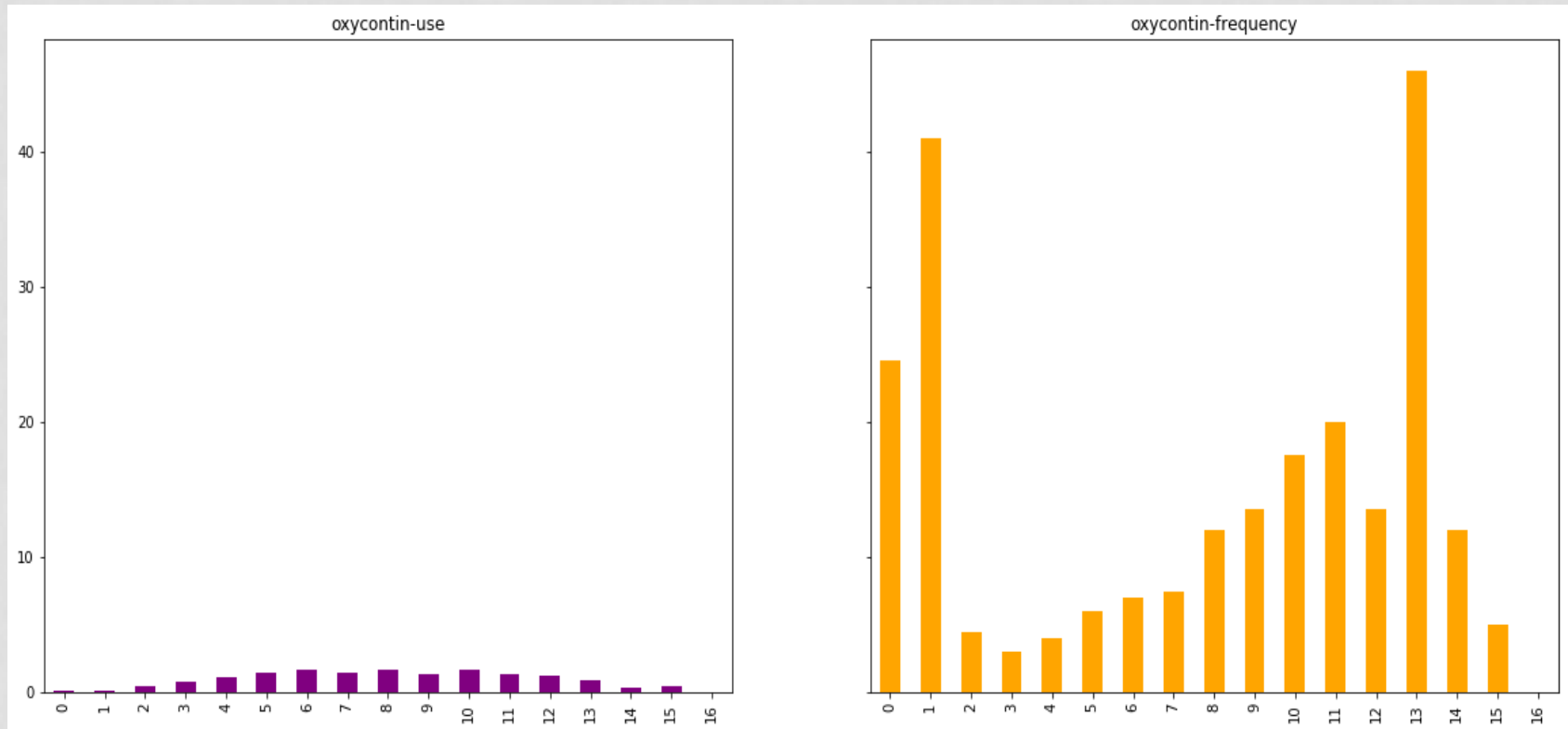
DRUGS AND THEIR FREQUENCY

- Pain reliever



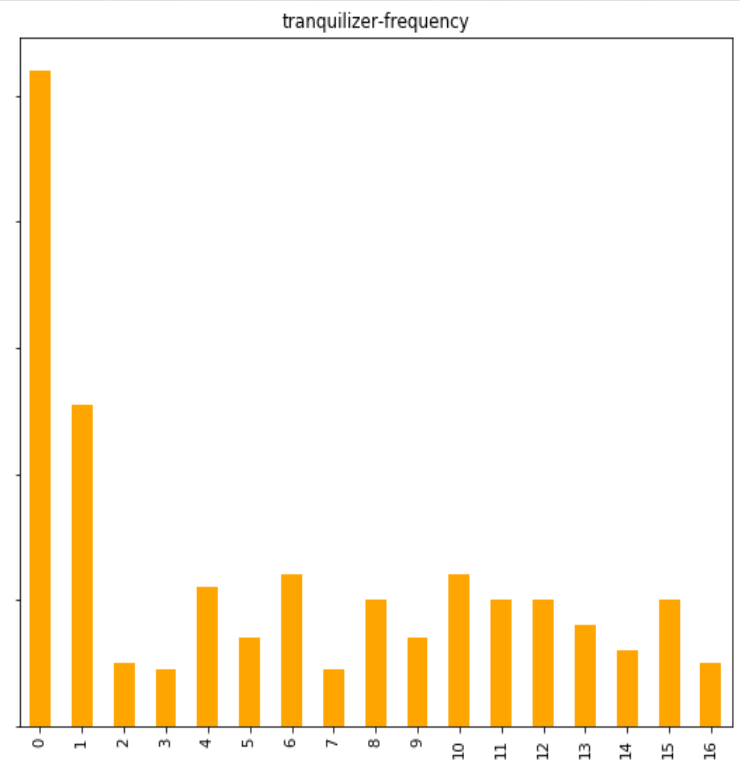
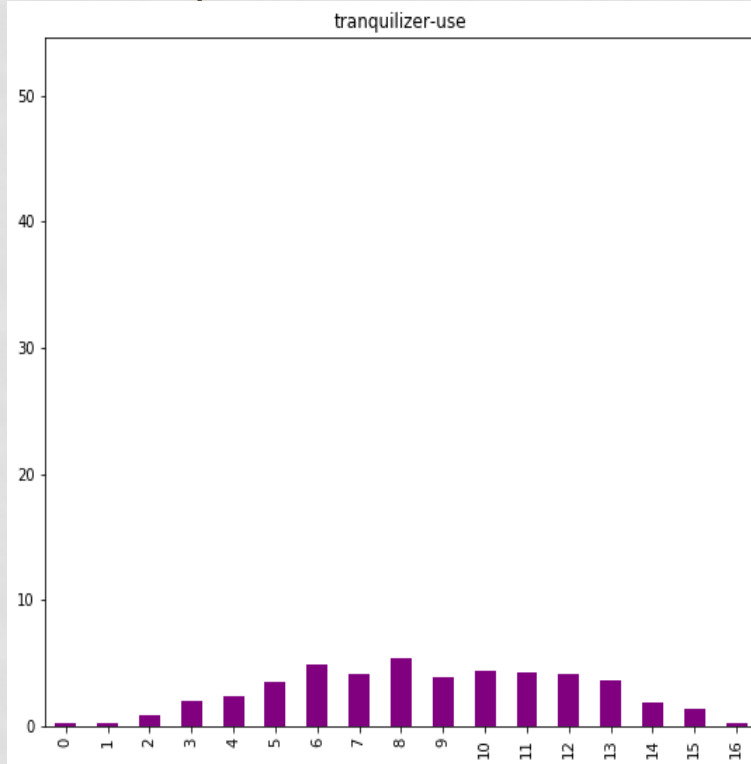
DRUGS AND THEIR FREQUENCY

- Oxycontin



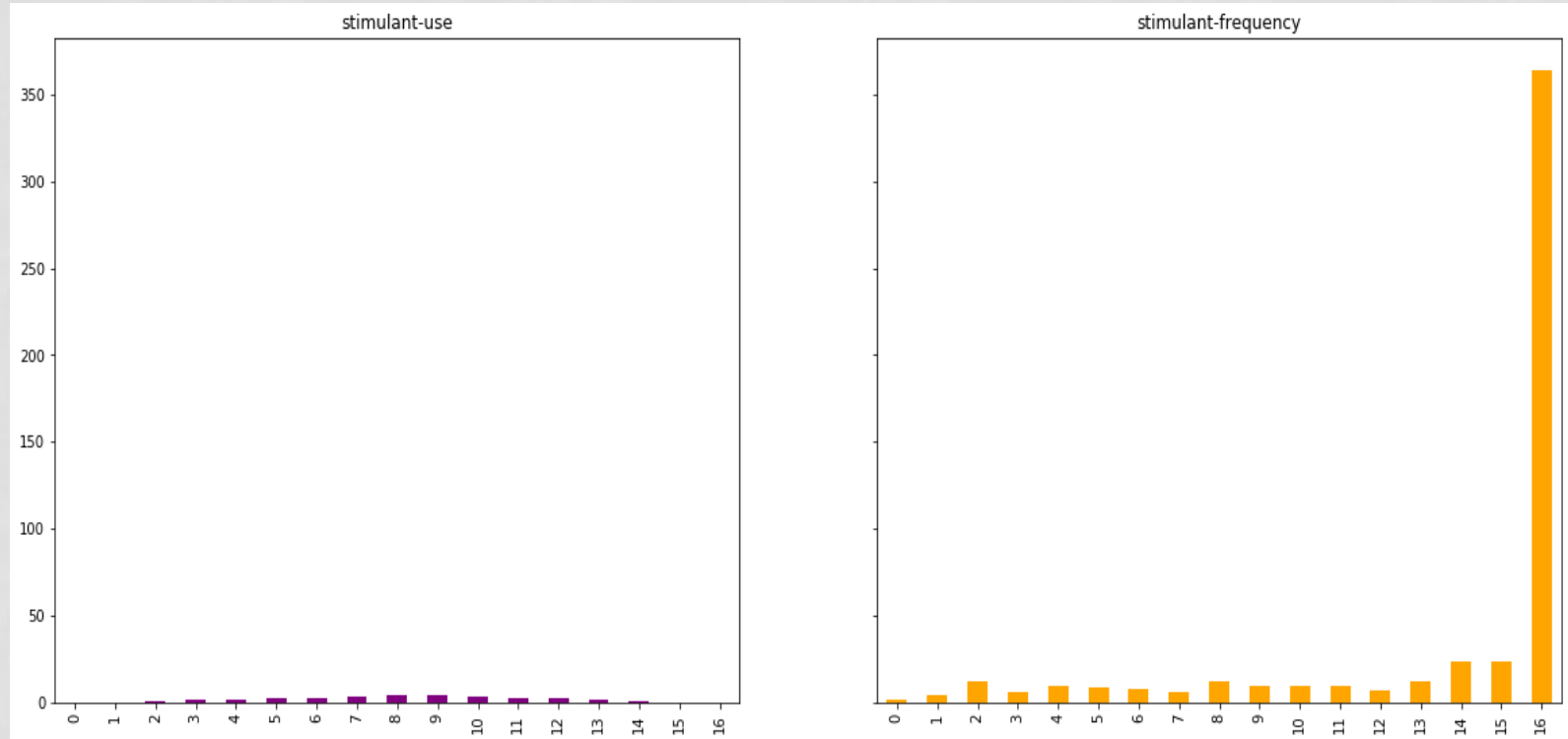
DRUGS AND THEIR FREQUENCY

- Tranquilizer



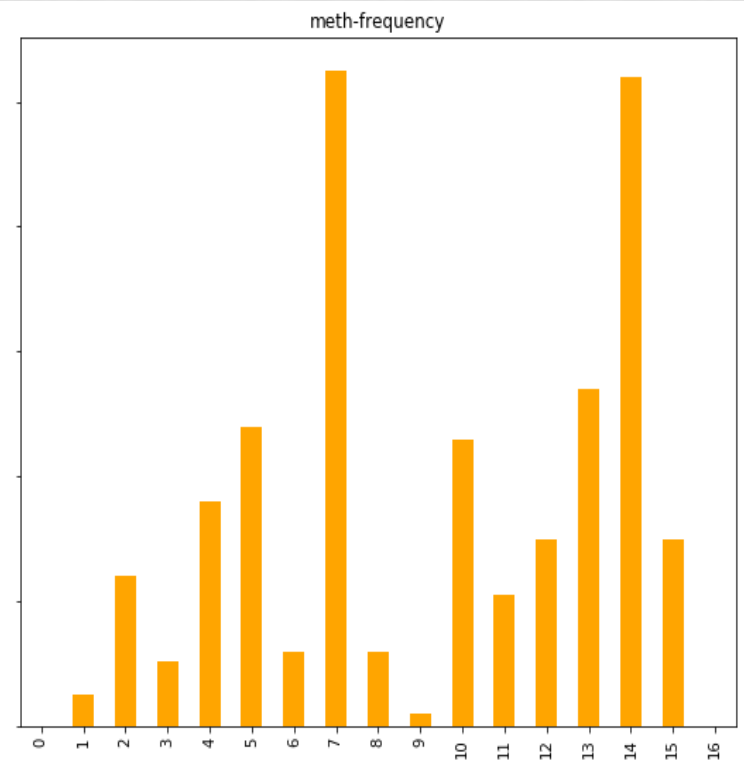
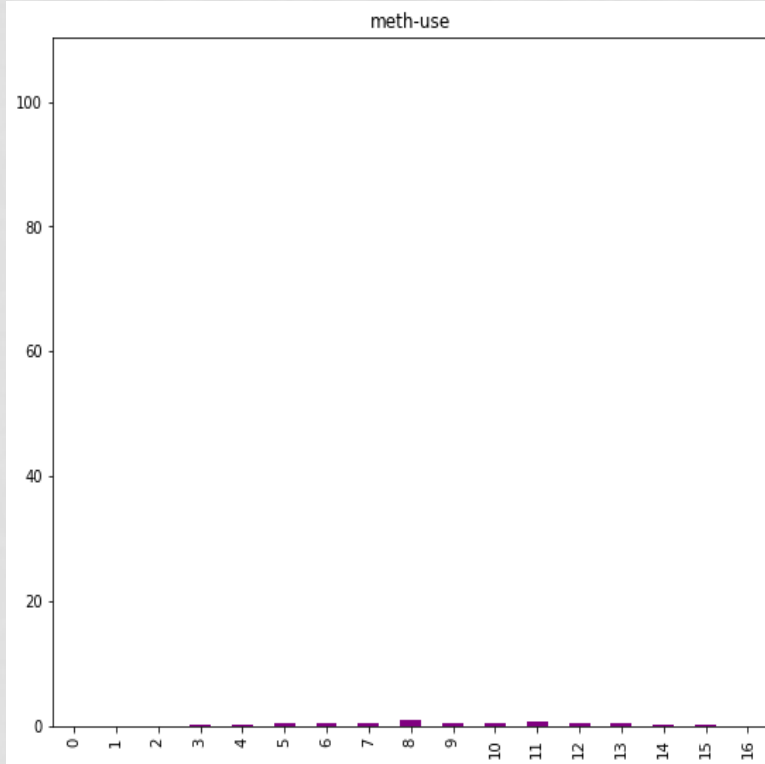
DRUGS AND THEIR FREQUENCY

- Stimulant



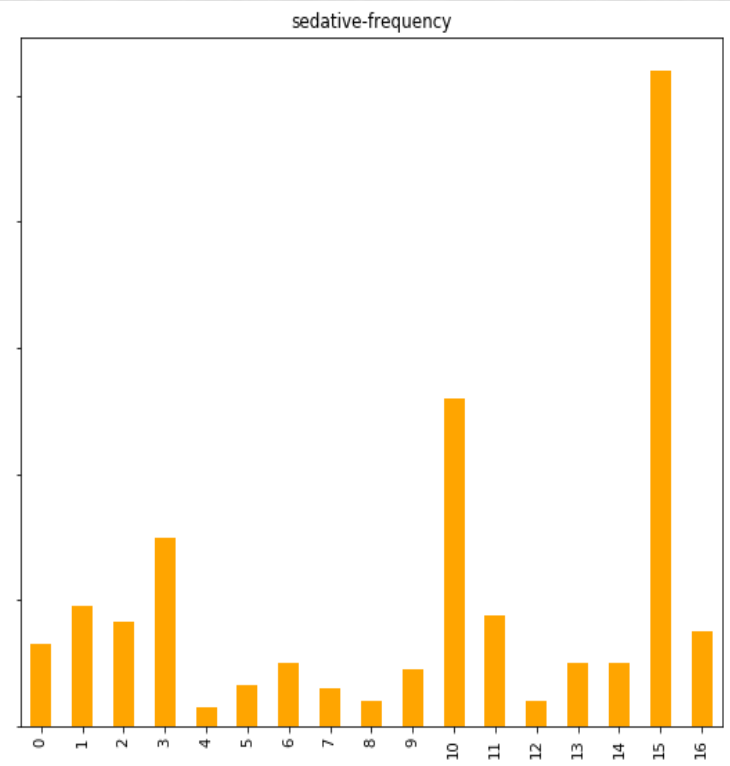
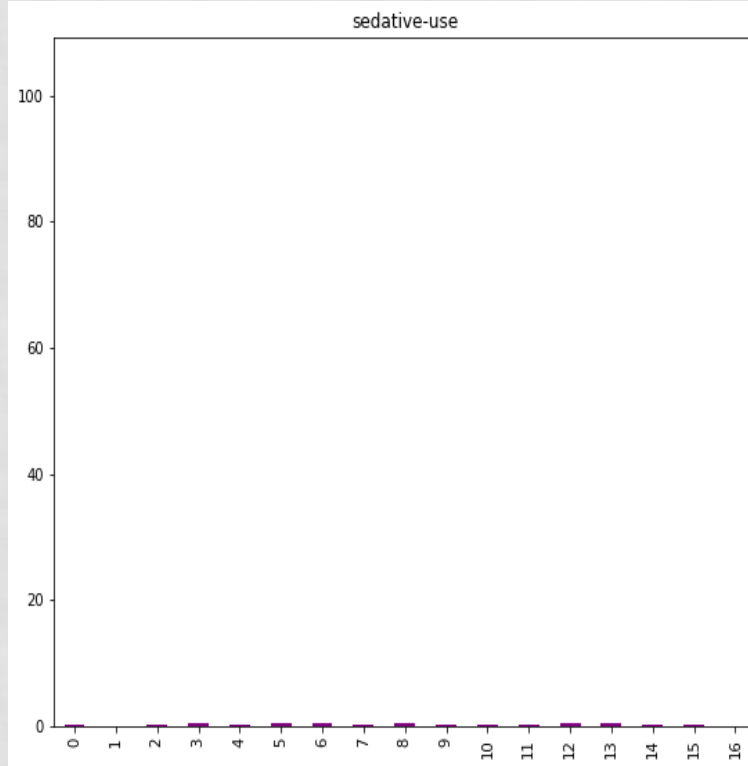
DRUGS AND THEIR FREQUENCY

- Meth



DRUGS AND THEIR FREQUENCY

- Sedative



CONCLUSION

- Findings show that needle related drugs abuses by numbers appear to be high around age groups between 18-23 age groups.
- Pain relievers abuse is among the highest and its abuse span across various age groups.
- There is also a strong correlation between tranquilizers with oxycontin and pain reliever. The correlations are really high at 0.96 between oxycontin and tranquilizer, and 0.98 between oxycontin and pain reliever; likely because tranquilizer has similar effects.

CONCLUSION

- Heroin has the highest frequency of abuse. This seems to imply that it is the most addictive needle related drug.
- Meth use is next in line.
- Hence, the monitoring of both drugs should be done with special attention for age groups of “19” and “35-39”.