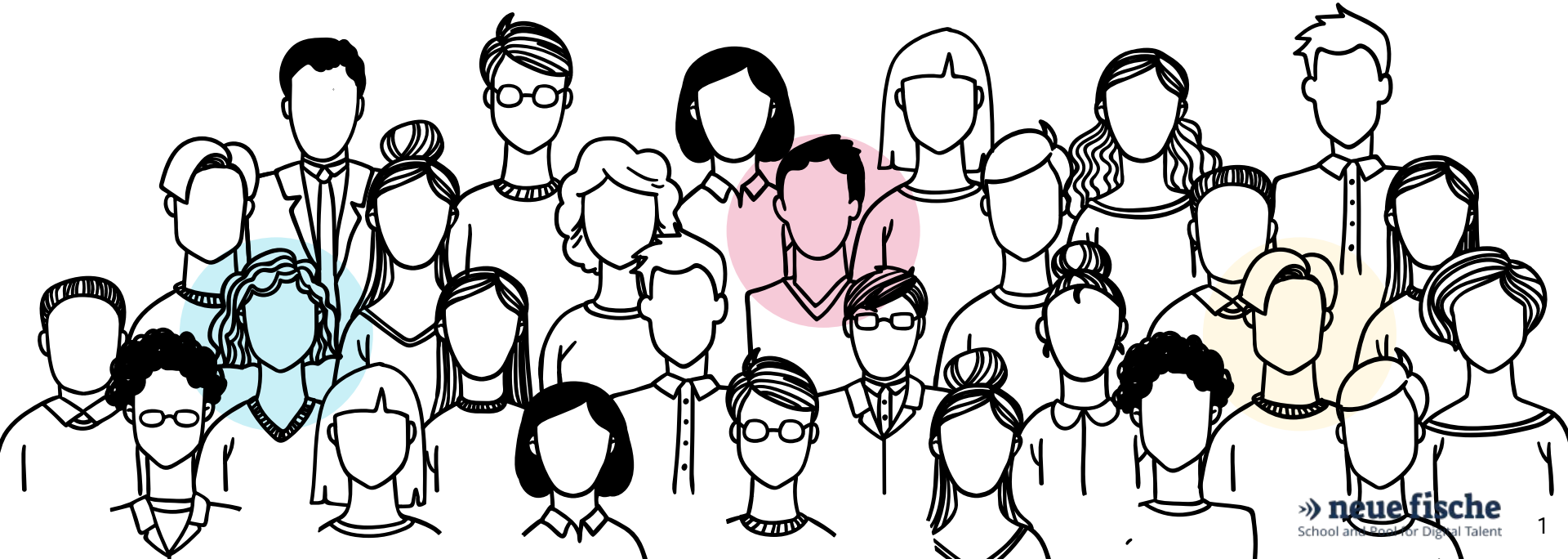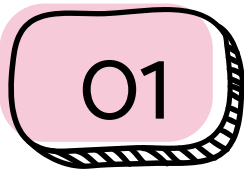# Personality and Drug Risk

By Matthis Westermann und Thomas Telgen
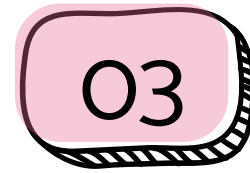
01

INTRODUCTION

What is it all about?

02

DATA SET

What are we working with?
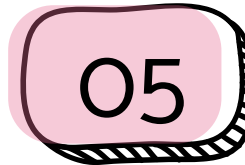
03

USER PREDICTION
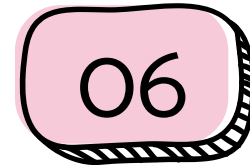
Who uses cannabis?

04
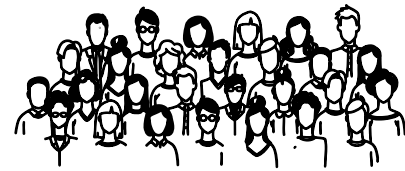
RESTRICTIONS

What did not work?

05

PERSONAL PREDICTION

What substances should you be careful with?
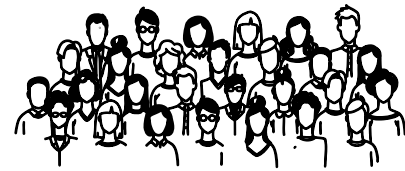
06

FUTURE WORK

What are the next steps?

01

INTRODUCTION

Correlation between personality and consumption habits?

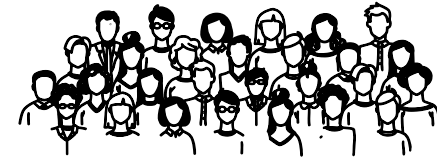Let's see what we can find in the Drug Consumption Data Set!

O2

DATA SET

What are we working with?
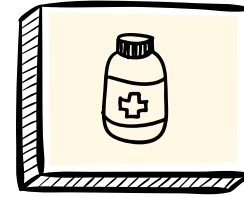
SOME OVERVIEW

1885 Participants

32 Features including:
- Age
- Education Level
- 7 Personal Traits
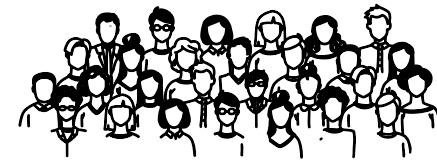- Use frequency for 18 Drugs
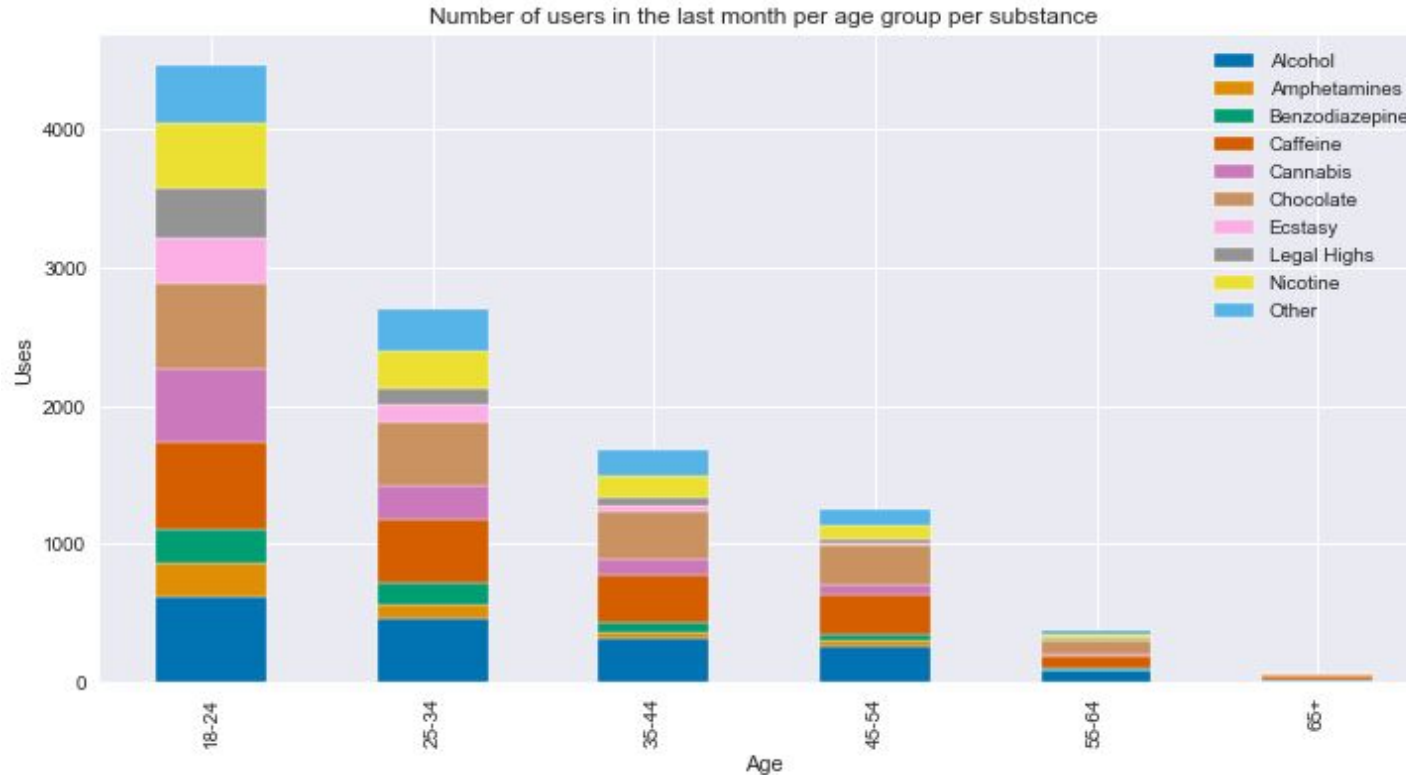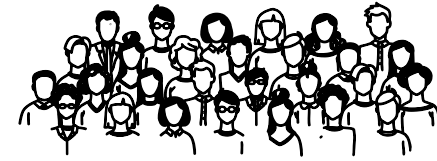- One made up Drug

Gender

Female	942
Male	943

Source

- Online Survey
- Mostly form UK and USA
- From 3/2011 to 3/2012

USE FREQUENCY BY AGE

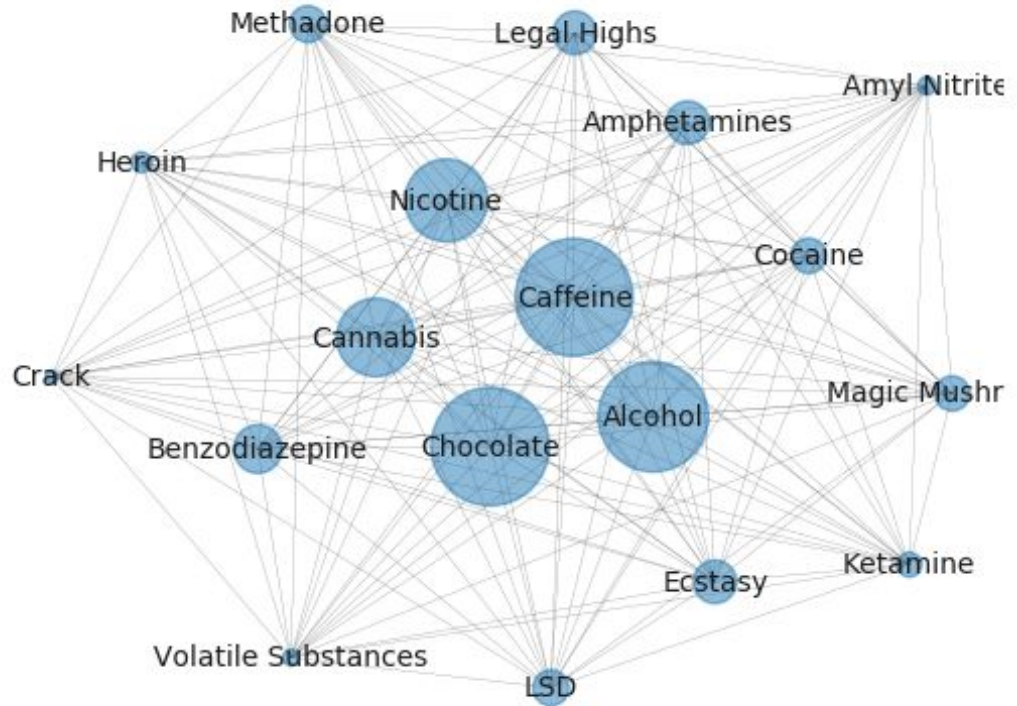Number of users in the last month per age group per substance
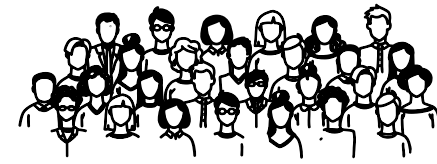
SUBSTANCE FRIENDS?

Findings:
- Four legal substances tend to cluster with Cannabis
- Other patterns not stable
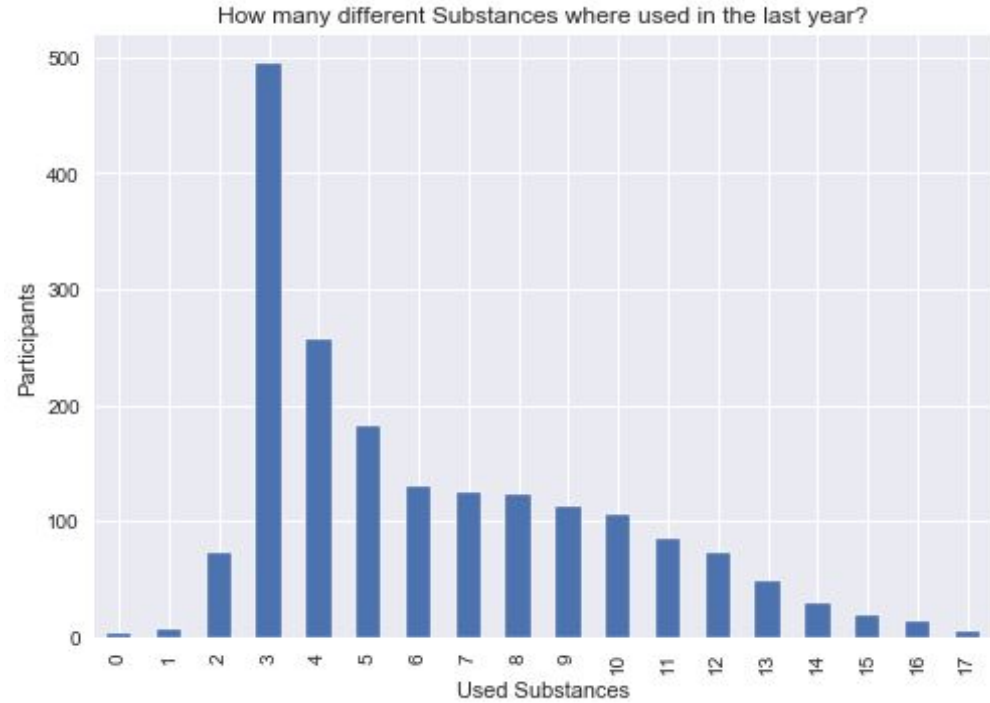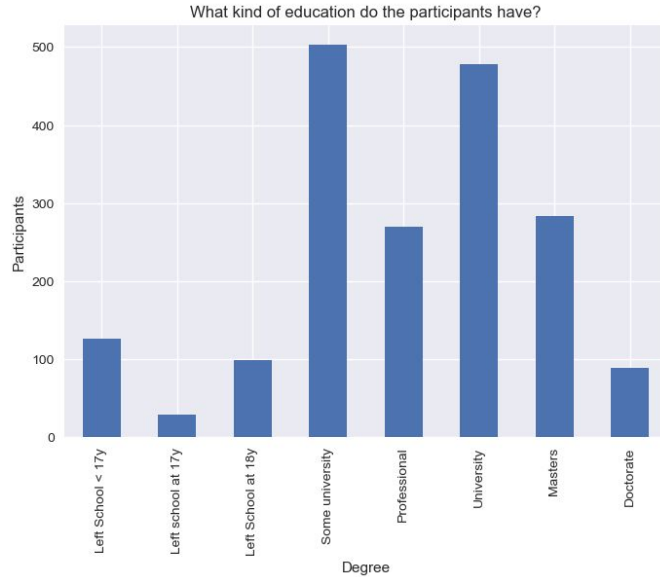- Cannabis and Ecstasy are the most linked illegal substances

Explanation of graph:
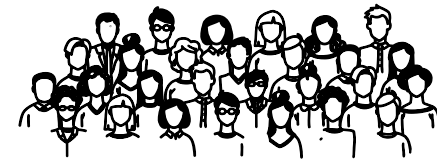More commonly used substances have larger nodes.
Substances more regularly used together are closer to each other.

WHAT ELSE?

What kind of education do the participants have?

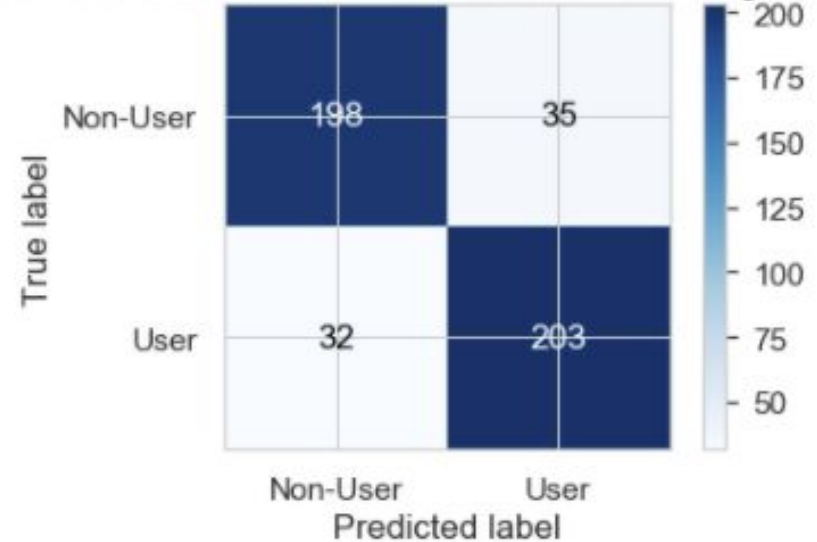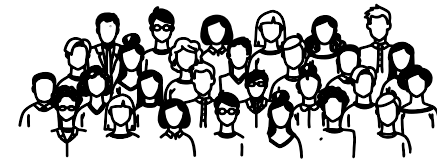How many different Substances where used in the last year?

03 - SUBSTANCE USER PREDICTION

- Drug use in binary classes (uses drugs, doesn't use drugs)
- Many imbalanced datasets which makes it hard to predict these values

Best models → SVC, Adaboost, Gradient Boosting

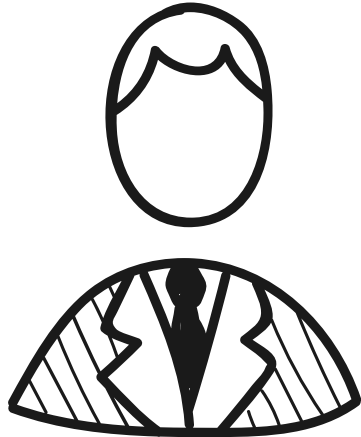Cannabis user confusion matrix: Gradient Boosting Classifier

04 - PREDICTIVE MODELING - RESTRICTIONS

- Drug use in multiple classes (ordinal values from 0 to 6)
- There were no predictions which were really good
    - Legal drugs → caffeine
    - Soft drugs → amyl nitrite, lsd
    - Hard drugs → cocaine, meth

- Amyl nitrite, lsd and vsa have a lighter imbalance than alcohol or crack
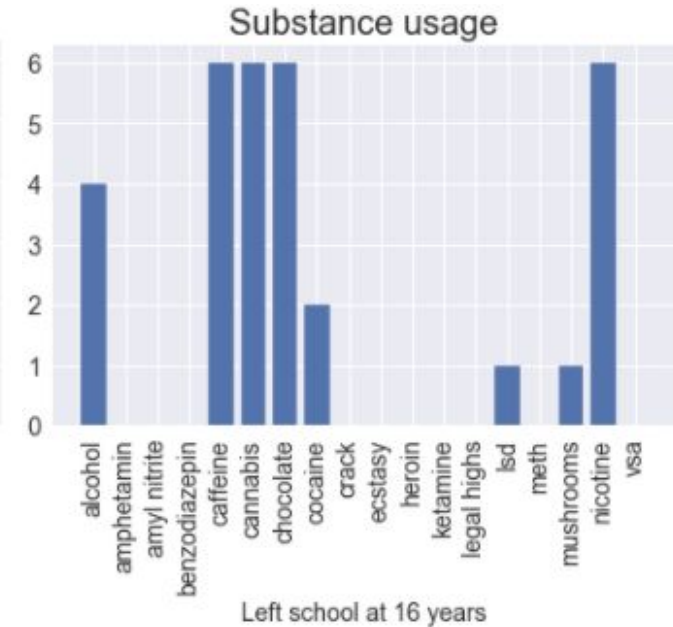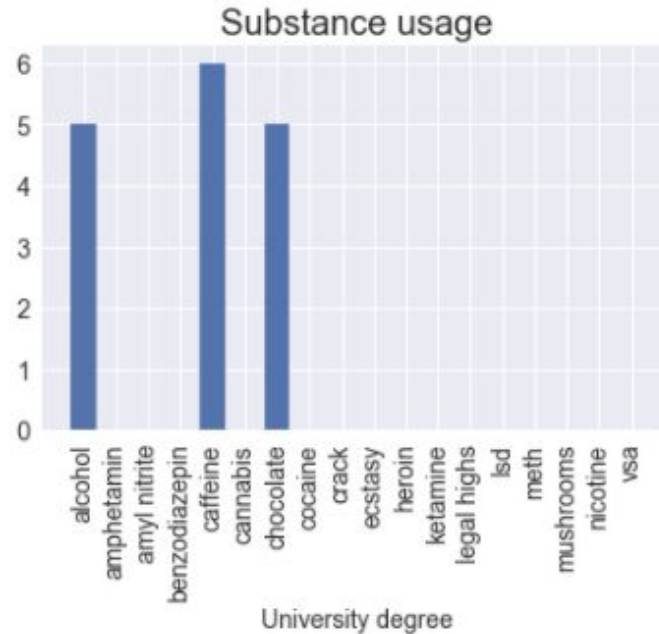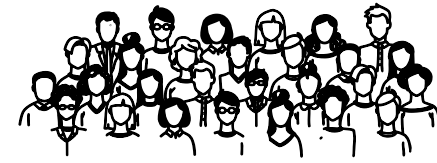- Can be predicted better by multiclass classification
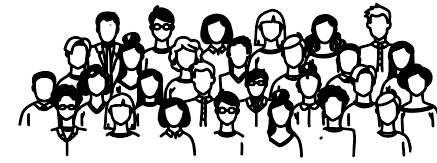
05 - WHAT SUBSTANCES FOR ME?

- Personal data
- Changing the educational background
- Model KNN

Test Subject A

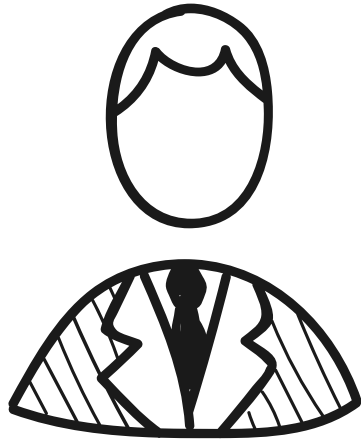Substance usage — University degree

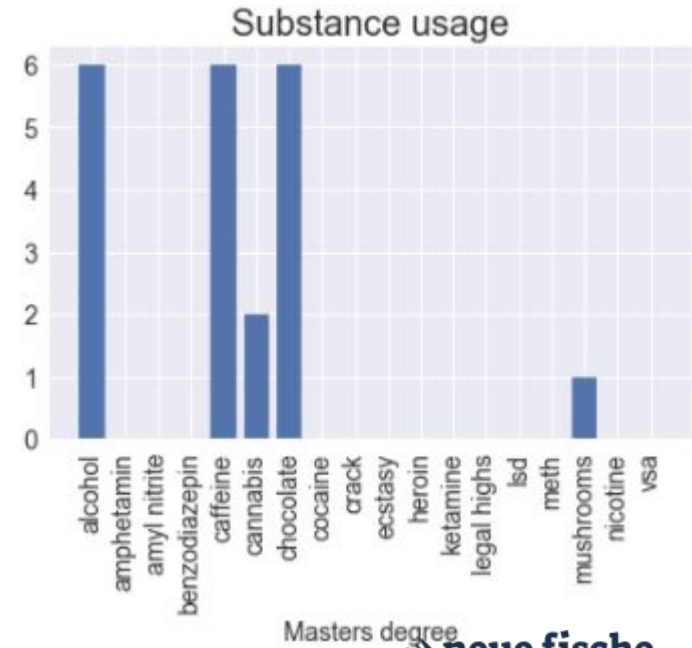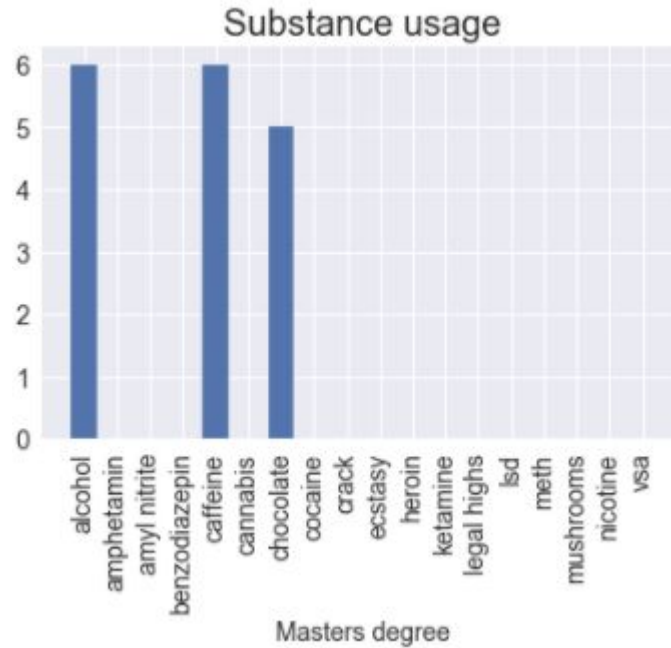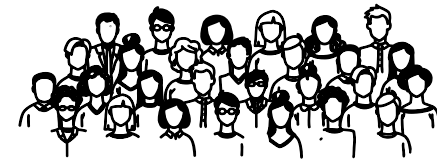Substance usage — Left school at 16 years

AND FOR HIM?

- Personal data
- Changing to higher impulsiveness and higher sensation seeking score
- Model KNN

Test Subject B

FUTURE WORK

What we'd like to try:
- Improving prediction by favoring substances with less occurrences (SMOTE)
- Focus on feature importance
- Developing an easy to use implementation for the substance "recommendation"

Questions to answer:
- What kind of substances are becoming less interesting over time?