

ANALISIS DATA "STUDENT ALCOHOL CONSUMPTION"

Menggunakan

EXPLORATORY DATA ANALYSIS (EDA)

CREDIT:

**Credit to kelompok kak Felix Aristo untuk
source code nya yang telah kami pakai
untuk project ini**

OLEH:

Kevin Raymond Metekohy (2019104778

M Attala R (2019104824)

M Dwi Ikhwannudin (2019104427)

M Alfath R (2019204766)

PROGRESS MINGGU KE-1

Pada minggu pertama kami melakukan proses anilisi pada dataset obesitas.

Latar Belakang

- Obesitas merupakan salah satu penyakit yang sering kita hiraukan. Dengan segala aspek kehidupan dapat kita lakukan secara digital, hal ini dapat membahayakan kesehatan kita masing-masing yang nantinya akan berpengaruh pada produktivitas kita.
- Berdasarkan dari data yang kami dapat, kami ingin mencari sebab yang membuat seseorang menjadi terjebak dalam pola hidup tidak sehat yang mengakibatkan Obesitas.

Pengertian

- Suatu gangguan yang melibatkan lemak tubuh berlebihan yang meningkatkan risiko masalah kesehatan.
- Obesitas sering kali terjadi karena kalori yang masuk lebih banyak daripada yang dibakar melalui olahraga dan kegiatan normal sehari-hari.
- Penanganan utamanya adalah perubahan gaya hidup seperti pola makan dan olahraga.

Manfaat

- Memberikan informasi mengenai insight yang bisa didapat khususnya terhadap wilayah yang terdampak dari dataset sehingga bisa diterapkan di kehidupan sehari-hari.
- Menjadi media pembelajaran kepada mahasiswa yang ingin mengetahui tentang analisis data.
- Memberikan informasi tentang penyebab terjadinya obesitas sehingga bisa bermanfaat untuk khalayak.

Tujuan

- Menemukan daerah terbesar penderita Obesitas
- Mencari sebab obesitas di daerah terbesar
- Menemukan range umur orang yang beresiko terkena obesitas

Data

- Obesity among adults by country, 1975–2016
- <https://www.kaggle.com/amanarora/obesity-among-adults-by-country-19752016>

PROGRESS MINGGU KE-2

Dikarenakan Pada minggu pertama datasets obesitas yang kami gunakan terlalu banyak fitur maka kami memutuskan mengganti dataset.

Latar Belakang

- Minuman beralkohol merupakan sejenis minuman yang sering dikonsumsi oleh manusia. Alkohol murni tidak dikonsumsi oleh manusia, manusia mengonsumsi minuman yang mengandung bahan sejenis alkohol, biasanya ethyl alcohol atau etanol. Namun belakangan ini alkohol sering kali disalahgunakan oleh pelajar terutama mereka yang masih dibawah umur
- Maka dari itu, kami disini akan mencoba mencari tahu sebab, dan akibat dari penyalahgunaan alkohol ini di kalangan pelajar sekolah menengah.

Pengertian

- Alkohol adalah jenis minuman yang mengandung unsur kimia etil alkohol atau etanol, etanol berbentuk cairan jernih, tidak berwarna dan memiliki rasa yang pahit.
- Alkohol adalah istilah yang sering dipakai untuk menyebut etanol, yang sering disebut juga “Grain alcohol” dan kadang minuman yang mengandung alkohol.
- Alkohol adalah minuman yang mengandung zat adiktif yang memiliki efek memabukkan dan merusak sistem saraf yang menyebabkan peminum akan merasakan kecanduan.

Manfaat

- Memberikan informasi mengenai insight yang bisa didapat khususnya terhadap wilayah yang terdampak dari dataset sehingga bisa diterapkan di kehidupan sehari-hari.
- Menjadi media pembelajaran kepada mahasiswa yang ingin mengetahui tentang analisis data.
- Memberikan informasi tentang penyebab terjadinya konsumsi alkohol di kalangan pelajar sehingga bisa bermanfaat untuk khalayak.

Tujuan

- Mencari alasan dan sebab yang membuat seorang pelajar untuk mengkonsumsi Alkohol
- Mencari akibat yang akan ditimbulkan dari konsumsi alcohol sejak dini.

Data

- Student Alcohol Consumption.

- <https://www.kaggle.com/uciml/student-alcohol-consumption>

Colab

- Analysis Data with Exploratory Data Analysis (EDA) Student Alcohol Consumption in Portugal
- <https://colab.research.google.com/drive/1DgqO3d8cD17XSixSarEJXL8pZjFaMjZT#scrollTo=qPUSeN0nuZxM>

Content

- School – student's school (binary: 'GP' – Gabriel Pereira or 'MS' – Mousinho da Silveira)
- Sex – student's sex (binary: 'F' – female or 'M' – male)
- Age – student's age (numeric: from 15 to 22)
- Address – student's home address type (binary: 'U' – urban or 'R' – rural)
- Famsize – family size (binary: 'LE3' – less or equal to 3 or 'GT3' – greater than 3)
- Pstatus – parent's cohabitation status (binary: 'T' – living together or 'A' – apart)
- Medu – mother's education (numeric: 0 – none, 1 – primary education (4th grade), 2 – 5th to 9th grade, 3 – secondary education or 4 – higher education)
- Fedu – father's education (numeric: 0 – none, 1 – primary education (4th grade), 2 – 5th to 9th grade, 3 – secondary education or 4 – higher education)

Content

- Mjob - mother's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at_home' or 'other')
- Fjob - father's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at_home' or 'other')
- Reason - reason to choose this school (nominal: close to 'home', school 'reputation', 'course' preference or 'other')
- Guardian - student's guardian (nominal: 'mother', 'father' or 'other')
- Traveltime - home to school travel time (numeric: 1 - <15 min., 2 - 15 to 30 min., 3 - 30 min. to 1 hour, or 4 - >1 hour)
- Studytime - weekly study time (numeric: 1 - <2 hours, 2 - 2 to 5 hours, 3 - 5 to 10 hours, or 4 - >10 hours)
- Failures - number of past class failures (numeric: n if $1 \leq n < 3$, else 4)

Content

- Schoolsup – extra educational support (binary: yes or no)
- Famsup – family educational support (binary: yes or no)
- Paid – extra paid classes within the course subject (Math or Portuguese) (binary: yes or no)
- Activities – extra-curricular activities (binary: yes or no)
- Nursery – attended nursery school (binary: yes or no)
- Higher – wants to take higher education (binary: yes or no)
- Internet – Internet access at home (binary: yes or no)
- Romantic – with a romantic relationship (binary: yes or no)
- Famrel – quality of family relationships (numeric: from 1 – very bad to 5 – excellent)
- Freetime – free time after school (numeric: from 1 – very low to 5 – very high)

Content

- Goout – going out with friends (numeric: from 1 – very low to 5 – very high)
- Dalc – workday alcohol consumption (numeric: from 1 – very low to 5 – very high)
- Walc – weekend alcohol consumption (numeric: from 1 – very low to 5 – very high)
- Health – current health status (numeric: from 1 – very bad to 5 – very good)
- Absences – number of school absences (numeric: from 0 to 93)

These grades are related with the course subject, Math or Portuguese:

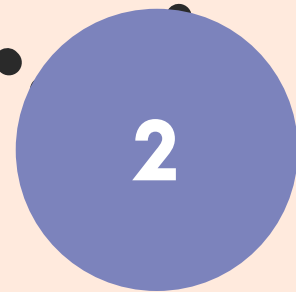
- G1 – first period grade (numeric: from 0 to 20)
- G2 – second period grade (numeric: from 0 to 20)
- G3 – final grade (numeric: from 0 to 20, output target)

Teknik

Data Acquisition

***Summary Data
Descriptive***

Data Visualization



Data Preprocessing

Data Clustering

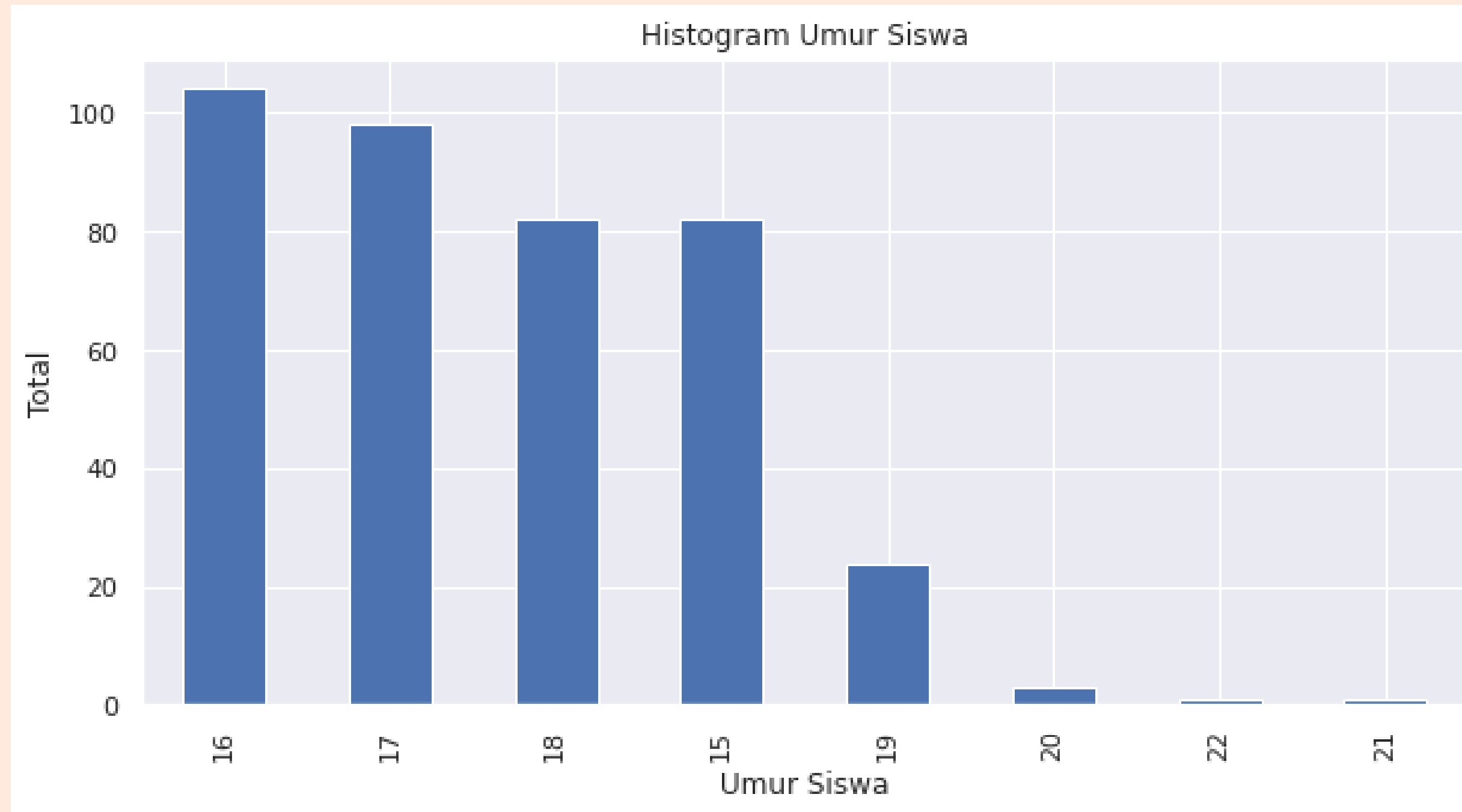
Artikel Pendukung

- 1.) <https://www.ijddr.in/drug-development/risk-factors-of-alcohol-consumption-among-portuguese-adolescents-and-young-adults-data-from-the-global-school-based-student-health-s.php?aid=7955>
- 2.) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5836530/>
- 3.) https://www.jstage.jst.go.jp/article/kenkokyoiku/19/2/19_135/_pdf
- 4.) <https://dspace.uevora.pt/rdpc/handle/10174/12467>
- 5.) <https://arrow.tudublin.ie/cgi/viewcontent.cgi?article=1103&context=scschcomdis>

PROGRESS MINGGU KE-3

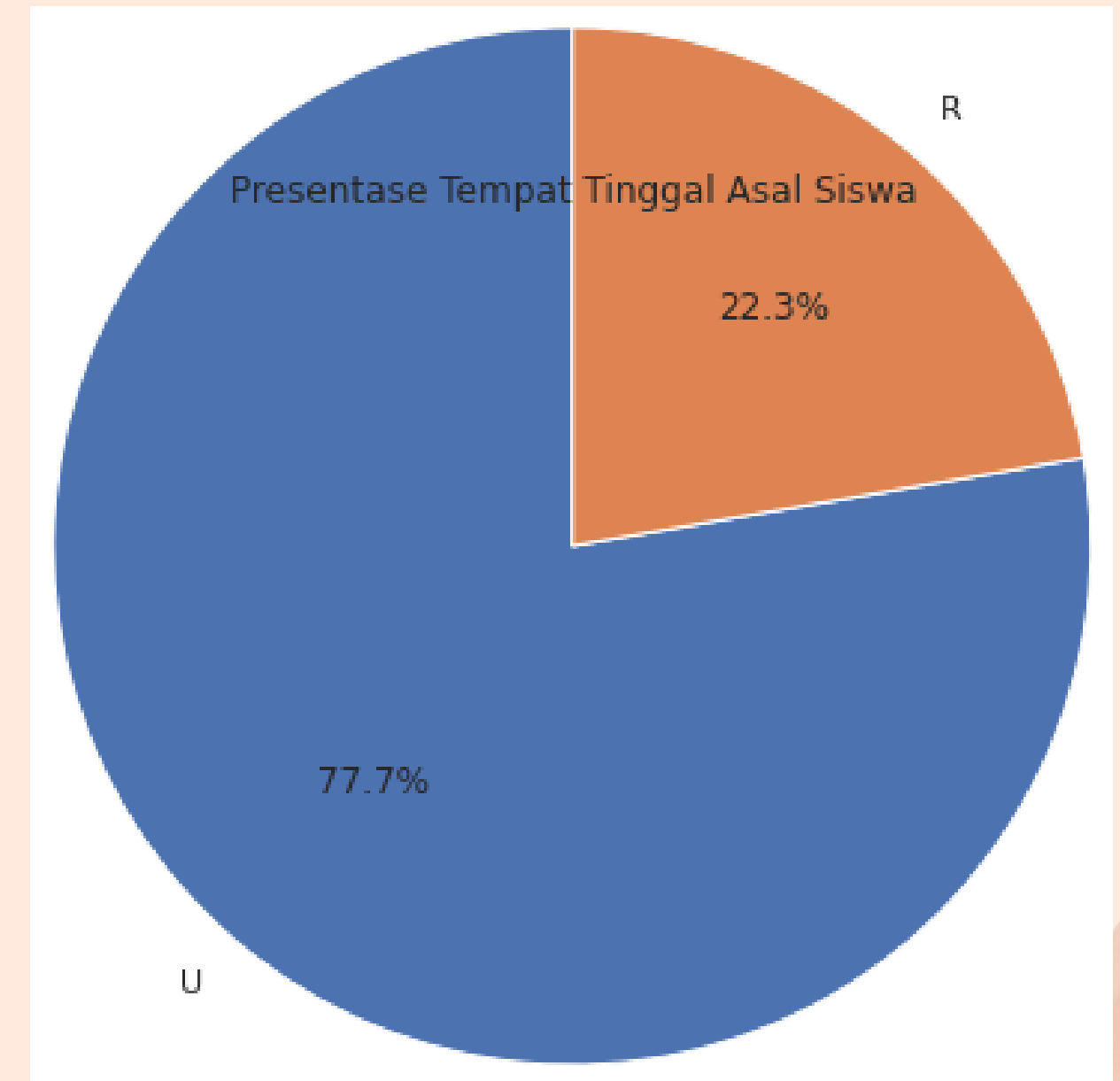
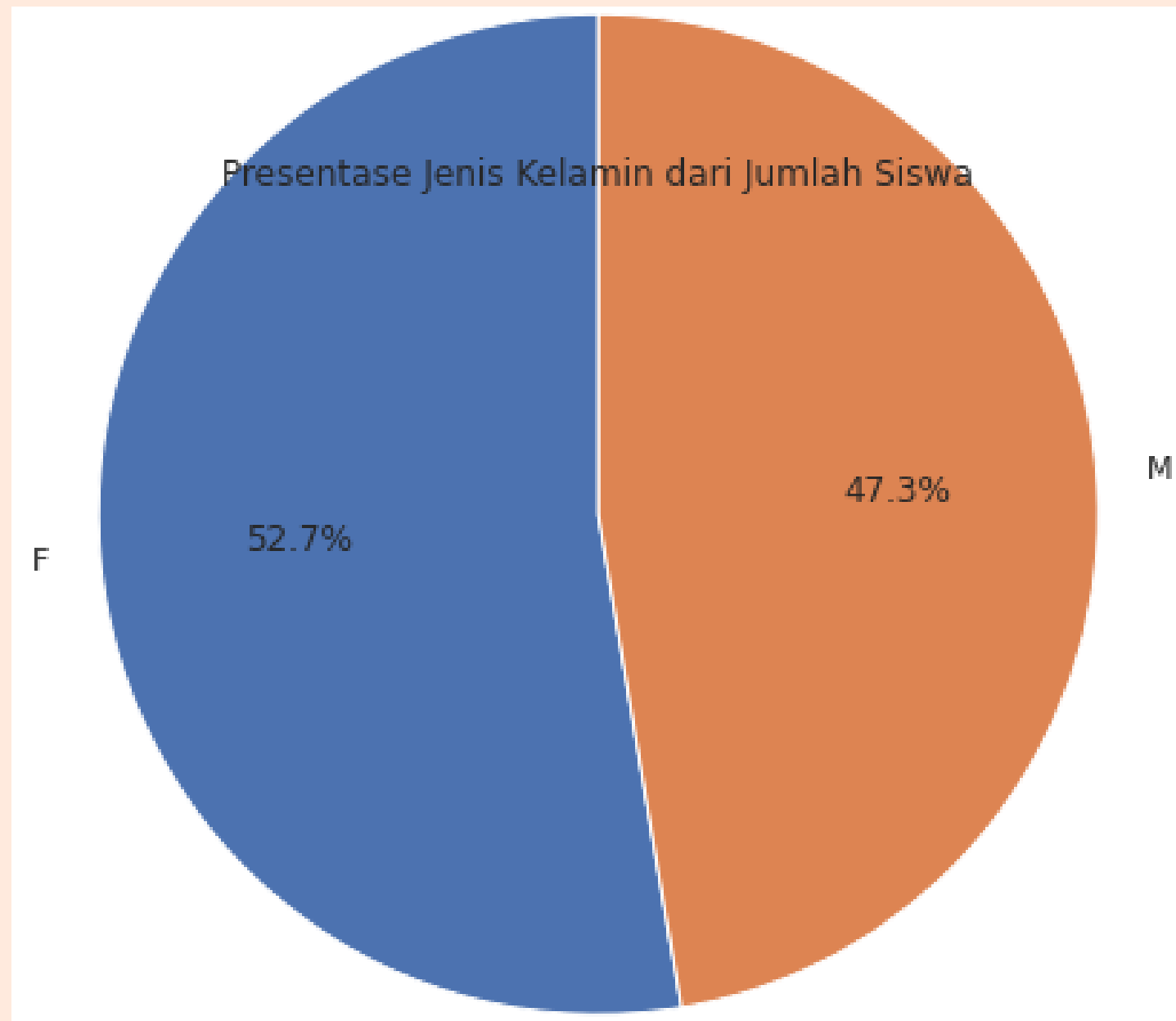
Pada minggu ketiga kami melakukan proses analisis data menggunakan beberapa library dan mendapat hasil visualisasi data. Dan kami melakukan proses clustering K-means namun karena diantara kelompok kami tidak ada yang bisa menjelaskan K-means maka kami memutuskan untuk tidak menampilkannya.

Visualisasi



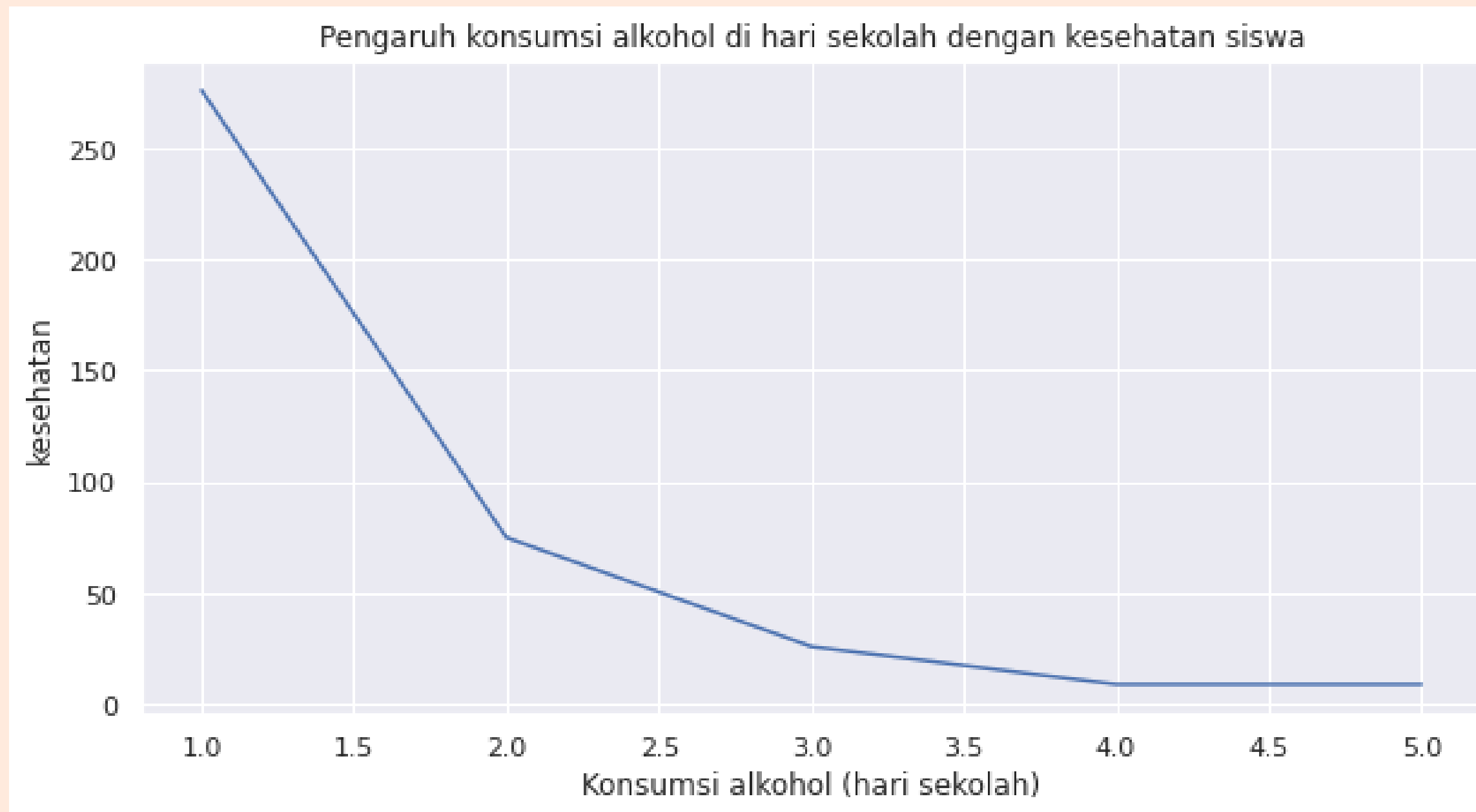
Subjek penelitian kami adalah sebuah sekolah di Portugal yang merupakan Sekolah Menengah Keatas. Mayoritas siswa disana masih berada di usia yang wajar yaitu di usia 15, 16, dan 17 tahun

Visualisasi



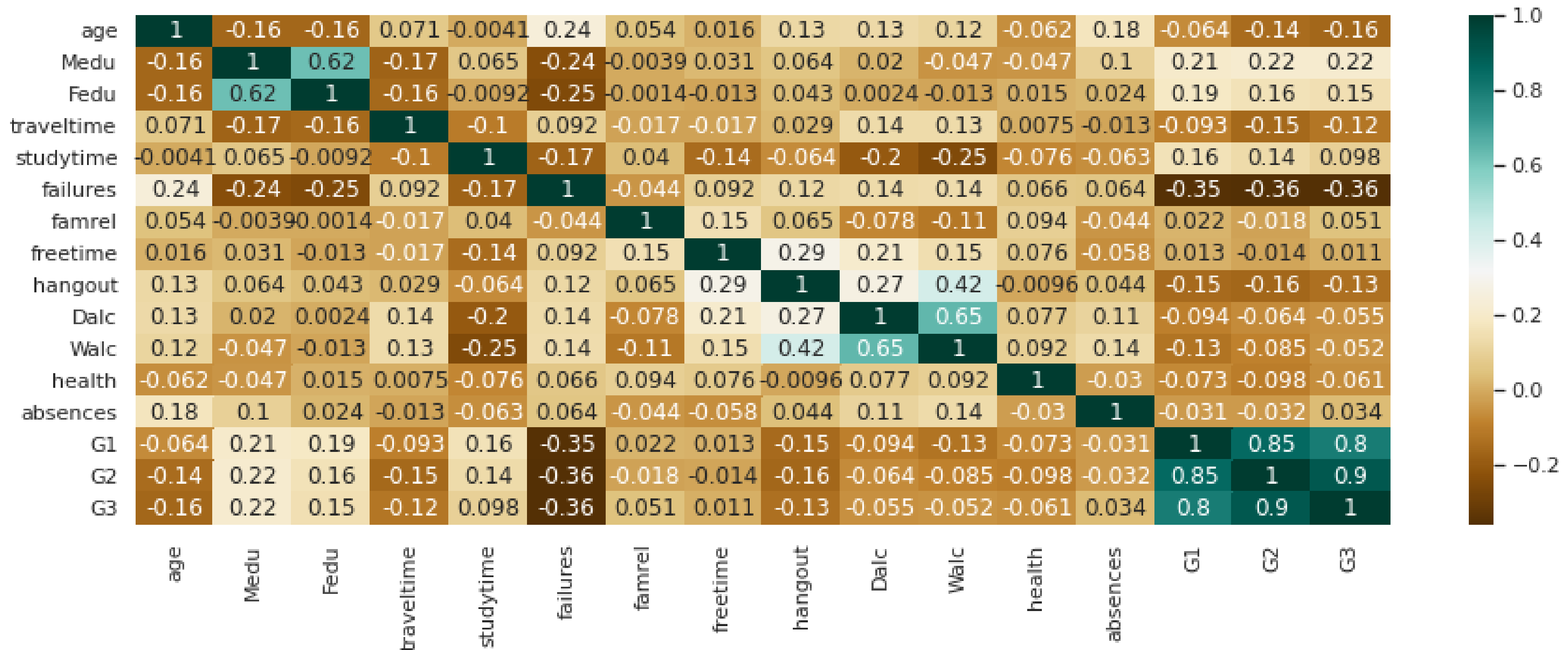
Subjek yang kami teliti terdiri dari subjek laki-laki dan perempuan dengan rasio yang kurang lebih seimbang. Lalu untuk tempat tinggal mereka rata-rata terdapat di area perkotaan (Urban) dengan beberapa ada yang tinggal di daerah kampung (Rural).

Visualisasi



Dapat dilihat terjadi degradasi kesehatan yang kontras ketika siswa tersebut memiliki tingkat konsumsi alkohol lebih dari sangat rendah. Hal ini sangat berpengaruh apalagi ketika dikonsumsi secara harian dapat mempengaruhi performa dan mental ketika berada di sekolah.

Visualisasi



Visualisasi

Selain faktor utama sebelumnya, kami menemukan beberapa faktor pendukung yang membuat seorang siswa dapat mendukung kesimpulan kami.

- Hubungan keluarga (famrel) dan tingkat nongkrong (hangout) seorang siswa memiliki keterkaitan dengan nilai 0.065 karena ketika seorang anak merasa mereka tidak cocok atau betah dengan keluarga mereka, mereka akan mencoba mencari keluarga diluar yang dapat berupa teman-teman mereka di sekolah.
- Tingkat nongkrong juga mempengaruhi konsumsi alkohol akhir pekan mereka dengan nilai 0.42 (hangout dengan Walc). Hal ini dikarenakan mereka menggunakan alkohol di tongkrongan mereka sebagai pelepasan dari penatnya hari sekolah ataupun keluarga mereka di rumah.
- Terakhir adalah tingkat konsumsi alkohol harian dan mingguan (Dalc dan Walc) memiliki nilai 0.65. Hal ini terjadi ketika mereka sudah nyaman dengan alkohol maka akan terjadi kecanduan sehingga mereka akan beranjak dari konsumsi yang sifatnya sekali-sekali menjadi sebuah rutinitas mereka yang selalu mereka lakukan setiap harinya.

PROGRESS MINGGU KE-4

Hasil analisa kami sudah kami buat dengan format booklet dan juga format report agar kami harap pembaca semakin memahami dan mengerti apa yang sudah kami analisa.

TERIMA KASIH!!