

Matson Social Skills

- Seyed Ehsan Sajadi
- Kian Jalilian
- Kamyab Abedi

DataMining Course - Spring 1400



TABLE OF CONTENTS

Introduction

Briefly elaborate on what we want to discuss.

PROBLEM & SOLUTION

Review on our datasets

OUR PROCESS

What we did?

TARGET

What is our target?

Our Data Set

Life Expectancy

The data was collected from WHO and United Nations website

Matson social skills

The data was collected by a master student of SBU

Fish

Review on Hematology factors collected by a PHD student of TU

Books Recommendation



Final Data Set

Matson social skills

The development of social skills is an important process in young childhood and adolescence. Deficits present in childhood that are left undetected and/or untreated can lead to increased problems into adulthood.

Life Expectancy

This study will focus on immunization factors, mortality factors, economic factors, social factors and other health related factors as well.



Matson social Skills



160 row

Contain 80 men and women



135 column

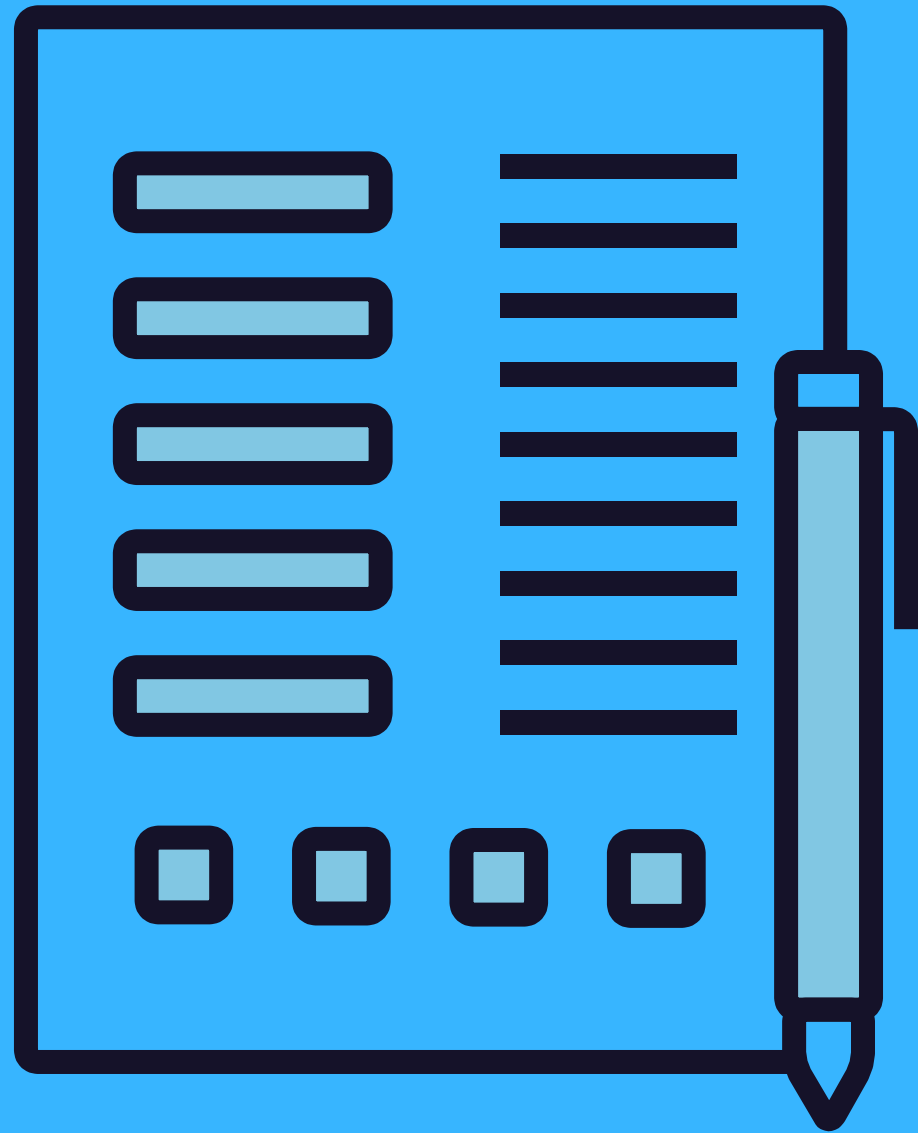
Collected by surveys



Final data set

Briefly elaborate on what you
want to discuss.

OUR PROCESS



1. Data Collection

The process of gathering information from different sources

2. Data Profiling

The process of reviewing source data for content and quality

سوالات مربوطه	بعد
1-18	مهارت های اجتماعی مناسب
19-29	رفتارهای غیر اجتماعی
30-41	پرخاشگری و رفتارهای تکانشی
42-47	برتری طلبی، اطمینان زیاد به خود داشتن
48-56	رابطه با همسالان

14	من به هنگام صحبت با دیگران، از آنها سوال هم می پرسم.	هرگز	به ندرت	برخی اوقات	اغلب اوقات	همیشه
15	من وقتی کسی را اذیت می کنم، بعدا متاسف می شوم.	هرگز	به ندرت	برخی اوقات	اغلب اوقات	همیشه
16	من به کسانی که با من خوب هستند، خوبی می کنم.	هرگز	به ندرت	برخی اوقات	اغلب اوقات	همیشه
17	من از اطرافیان احوالپرسی کرده و جویای کار آنها و مسایلی از این قبیل می شوم.	هرگز	به ندرت	برخی اوقات	اغلب اوقات	همیشه

بهزیستی روانشناختی

13-25-28-29-48-59-65-73	رشد و بالندگی فردی
1-11-20-24-37-47-58-69-71-74	خود پیروی
5-7-9-10-21-40-41-64	ارتباط مثبت با دیگران
4-12-19-22-23-30-31-33-34-38-45-49-52-55-57-60-61-62- 68	رضایت از زندگی
3-8-14-16-27-32-39-42-46-50-56-63-76	معنویت
2-6-15-17-18-26-35-36-43-44-51-53-54-66-67-70-72-75-77	شادی و خوشبینی

15	من جزء افرادی هستم که دوست دارند چیزهای جدید را تجربه کنند	کاملاً موافقم	موافقم	نظری ندارم	مخالقم	کاملاً مخالفم
16	معنی وجودی خودم را درک می کنم	موافقم	کاملاً موافقم	نظری	مخالقم	کاملاً
17	در کل، نگرش مثبت و خوشبینانه ای نسبت به انسان ها و زندگی دارم	کاملاً موافقم	موافقم	نظری ندارم	مخالقم	کاملاً مخالفم



3. Data Exploration

A visual exploration to help users familiarize with information

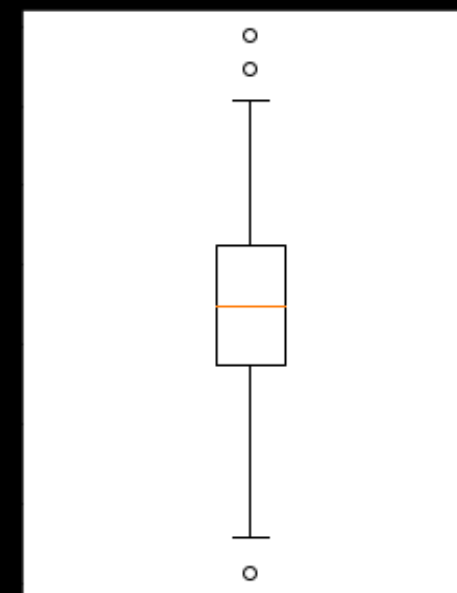
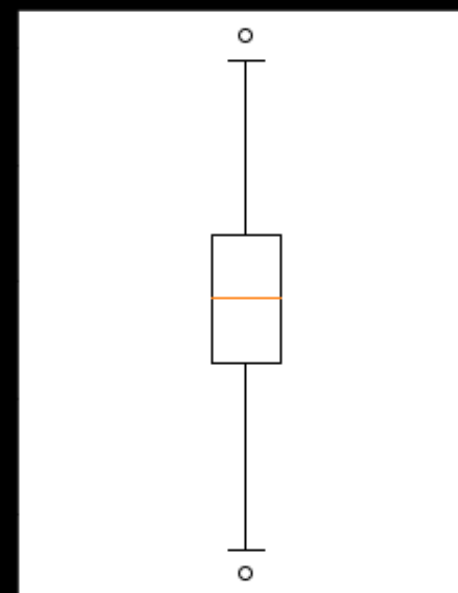
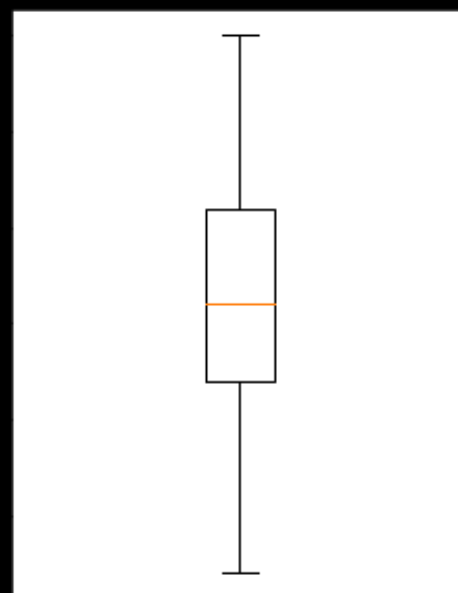
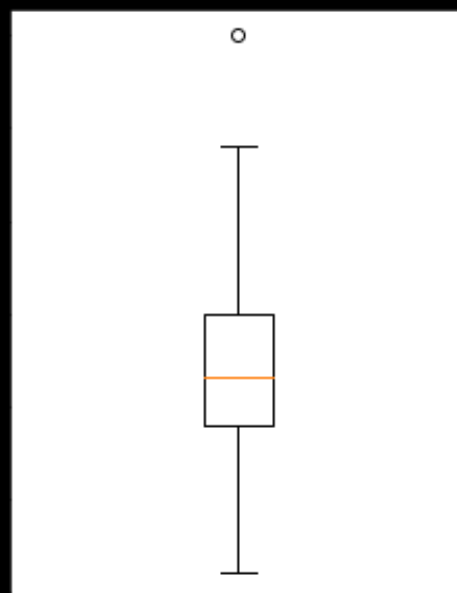
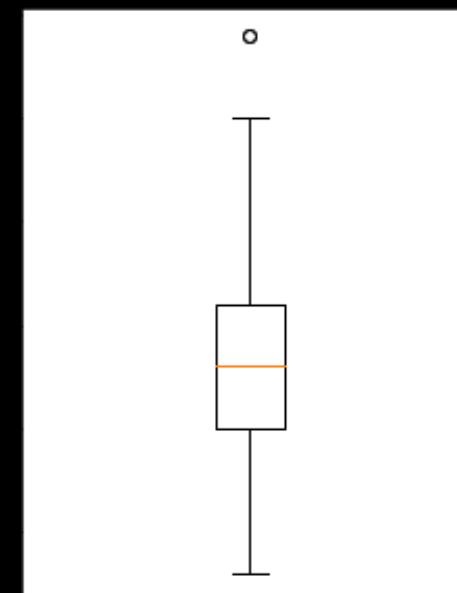
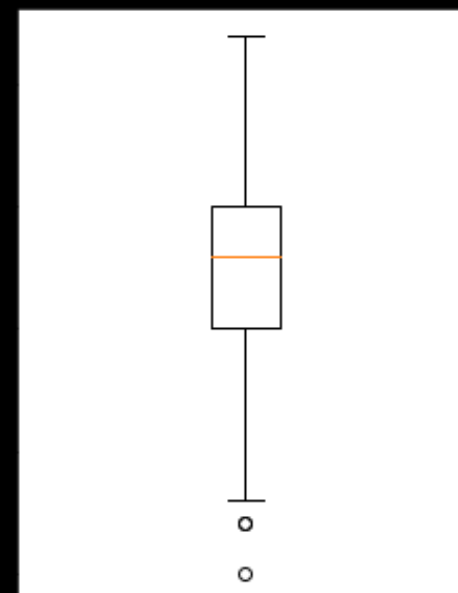
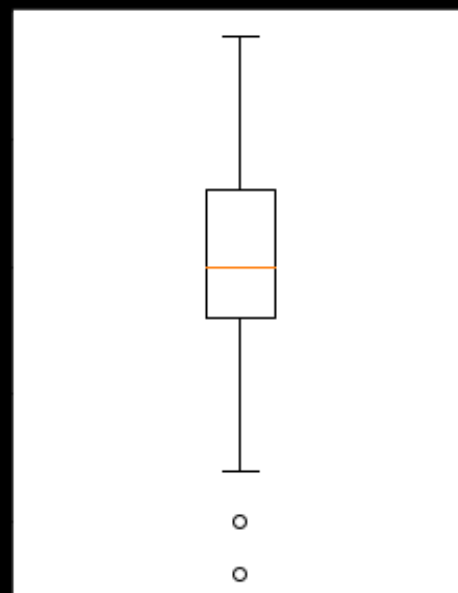
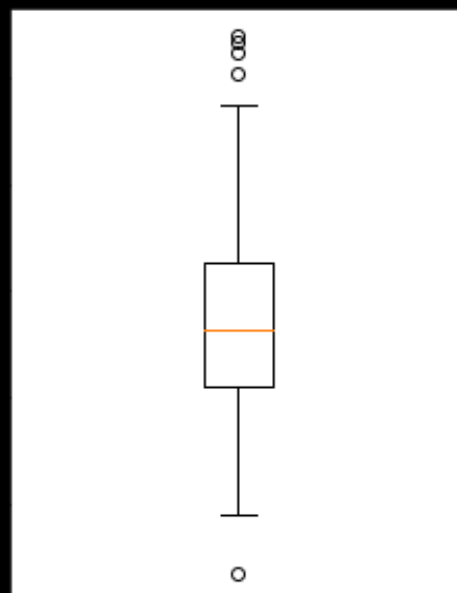
```
[155]: # Discard the metadata (age and variable and row).
le = le.drop(['age', 'variable', 'row'], axis=1)
```

```
[156]: le.head()
```

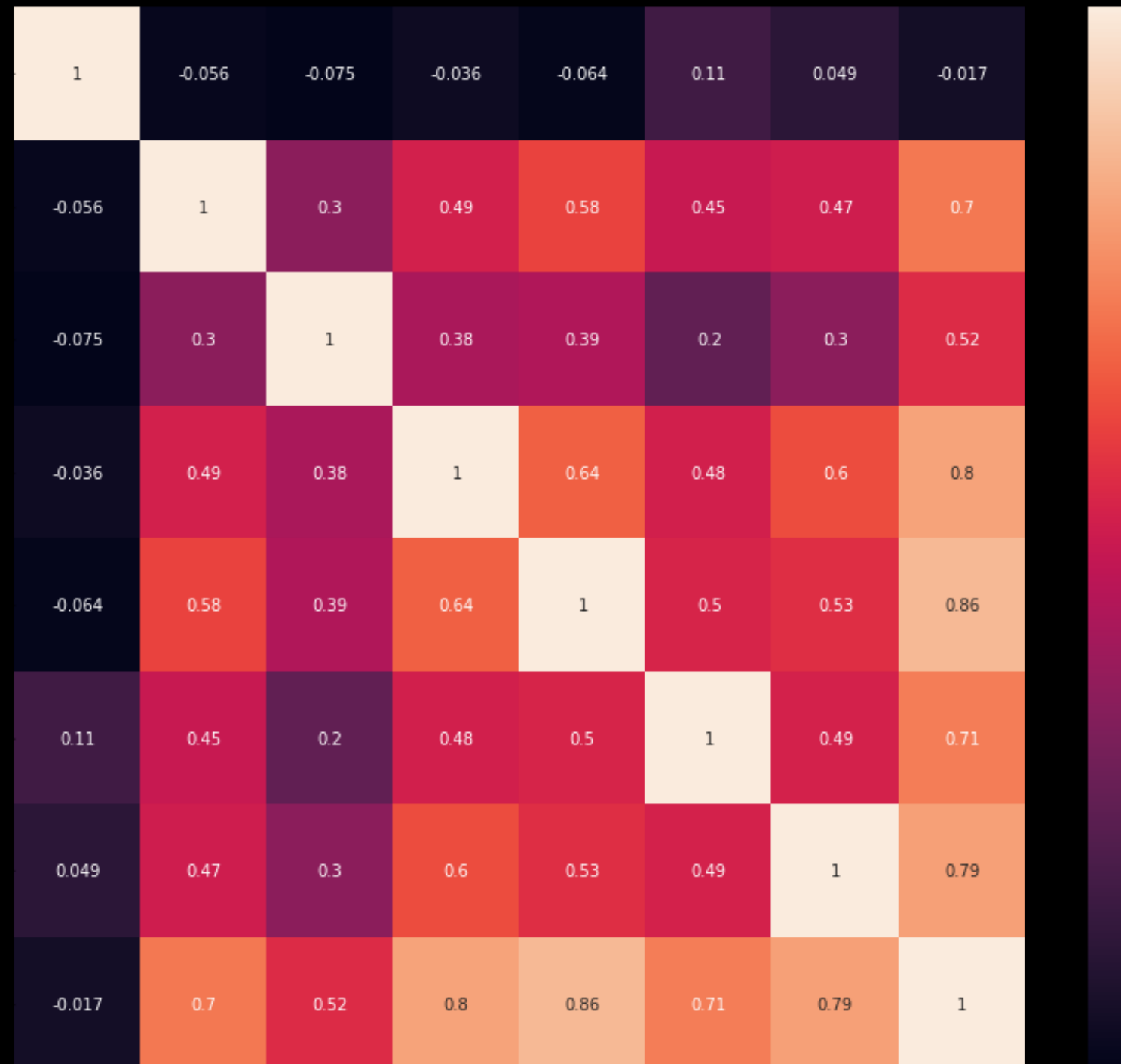
[156]:	sex	b1	b2	b3	b4	b5	b6	b7	b8	b9	...	m55	m56	maharat	roshd	khodpeiravi	ertebatm	rezayat	manaviat	shadi	behzisti
0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	...	1.0	5.0	168.0	15.0	28.0	16.0	42.0	22.0	48.0	170.0
1	1.0	3.0	1.0	2.0	1.0	2.0	1.0	1.0	1.0	1.0	...	5.0	5.0	170.0	18.0	26.0	17.0	43.0	21.0	47.0	171.0
2	1.0	3.0	4.0	4.0	1.0	1.0	1.0	1.0	1.0	1.0	...	1.0	1.0	135.0	15.0	31.0	18.0	34.0	27.0	42.0	171.0
3	1.0	2.0	5.0	2.0	2.0	2.0	2.0	1.0	2.0	2.0	...	1.0	1.0	155.0	24.0	29.0	20.0	45.0	25.0	53.0	193.0
4	1.0	1.0	1.0	2.0	1.0	1.0	1.0	1.0	1.0	1.0	...	1.0	1.0	152.0	24.0	25.0	12.0	40.0	22.0	38.0	159.0

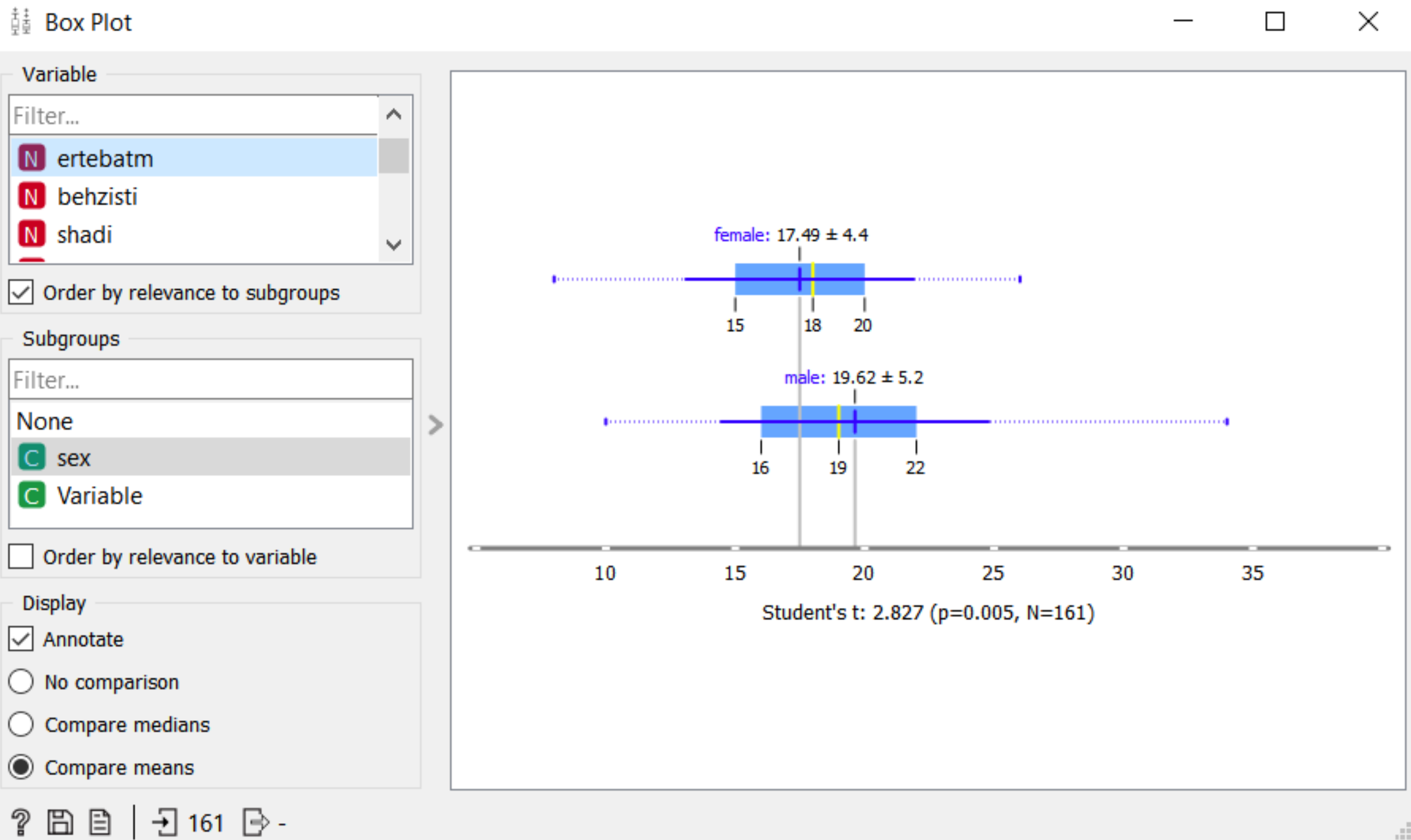
5 rows × 142 columns

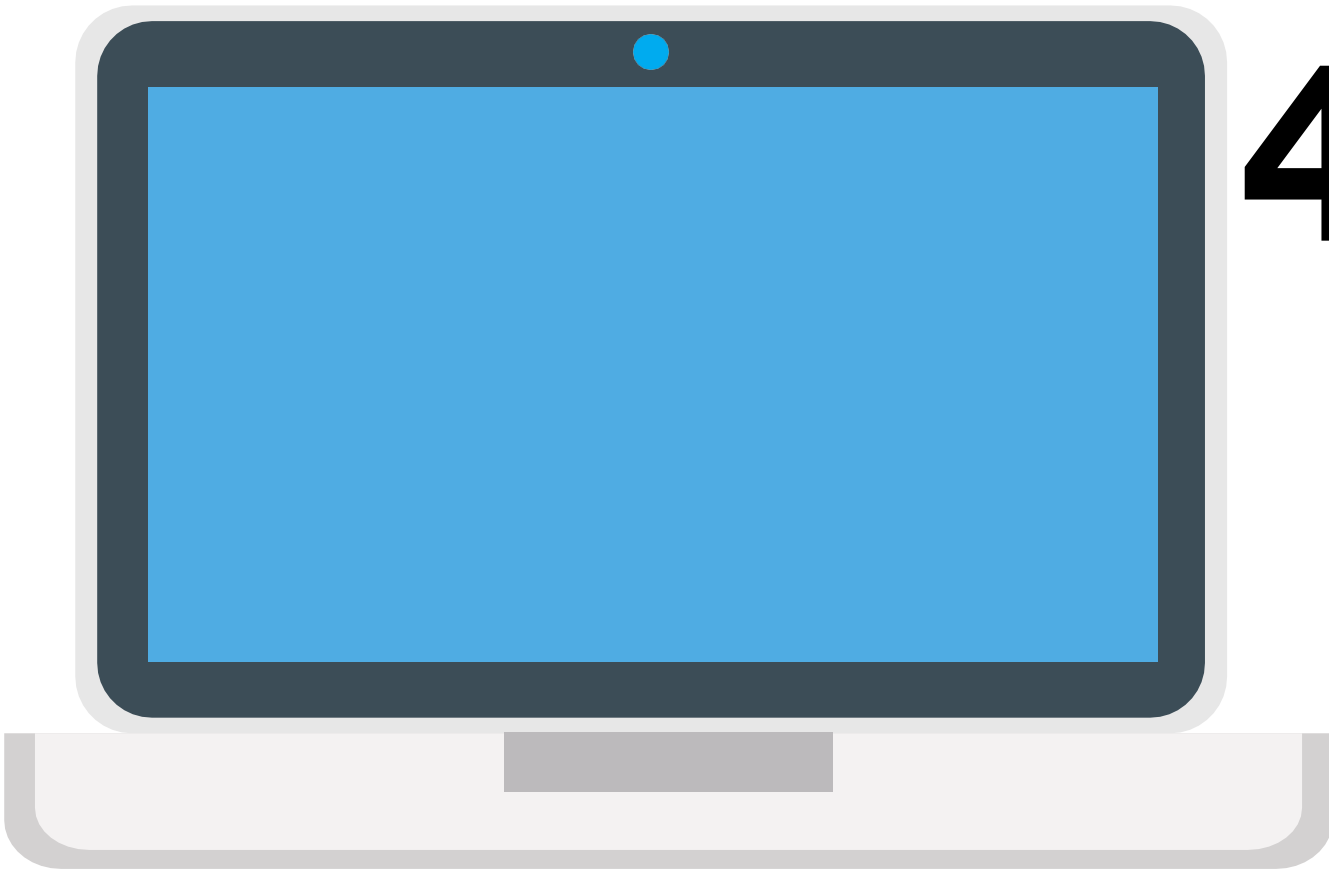
```
[163]: # Visualize the cleaned data for each feature using box plots.  
plt.figure(figsize=(18,30))  
  
for variable,i in col_dict.items():  
    plt.subplot(5,4,i)  
    plt.boxplot(le[variable],whis=1.5)  
    plt.title(variable)  
  
plt.show()
```



```
[164]: # Plot heatmap to visualize the correlations.  
plt.figure(figsize = (14, 12))  
sns.heatmap(le.corr(), annot = True)  
plt.title('Correlation between different features');
```







4. Pre-Processing

It include feature engineering, management of missing values and outliers, data normalization, etc.



OUR TARGET



in progress

Do you have any questions?

THANKS