

Stress_level_analysis

Trần Đăng Khôi

2025-06-19

Giới thiệu đề tài

Ở dataset 1 này, nhóm tập trung vào phân tích thống kê nhiều chiều nhằm khám phá các yếu tố ảnh hưởng đến mức độ căng thẳng của học sinh. Thông qua các kỹ thuật như phân tích thành phần chính (PCA), phân tích nhân tố (FA) và phân tích phương sai nhiều chiều (MANOVA), đề tài hướng đến việc rút gọn số lượng biến để dễ dàng phân tích hơn mà không có mất quá nhiều thông tin của bộ dữ liệu, phát hiện các cấu trúc tiềm ẩn trong dữ liệu để xem xét những nhân tố ẩn chứa sâu bên trong bộ dữ liệu, và cũng kiểm tra thử sự khác biệt giữa các nhóm đối tượng. Dữ liệu được sử dụng bao gồm nhiều biến liên quan đến các khía cạnh tâm lý, sinh lý, xã hội, môi trường và học tập – những yếu tố được cho là có mối liên hệ chặt chẽ với mức độ căng thẳng. Mục tiêu cuối cùng là đưa ra những nhận định có giá trị thực tiễn trong việc hiểu và cải thiện sức khỏe tinh thần cho học sinh trong môi trường học đường.

Giới thiệu dữ liệu

Bộ dữ liệu StressLevelDataset.csv là một tập dữ liệu nghiên cứu về mức độ căng thẳng của sinh viên, được trích xuất từ nguồn chia sẻ dữ liệu trực tuyến Kaggle. Dữ liệu bao gồm 1100 quan trắc và 21 biến, phản ánh toàn diện các yếu tố có thể ảnh hưởng đến trạng thái tâm lý của sinh viên. Dựa vào thông tin trên Kaggle, các biến trong dữ liệu được xây dựng dựa trên 5 nhóm yếu tố chính và bao gồm các biến: - Tâm lý: - Mức độ lo lắng (anxiety-level) - Lòng tự trọng (self-esteem) - Tiền sử sức khỏe tâm thần (mental-health-history) - Trầm cảm (depression)

- Sinh lý:
 - Đau đầu (headache)
 - Huyết áp (blood-pressure)
 - Chất lượng giấc ngủ (sleep-quality)
 - Vấn đề hô hấp (breathing-problem)
- Môi trường:
 - Mức độ ồn (noise-level)
 - Điều kiện sống (living-conditions)
 - Sự an toàn (safety)
 - Nhu cầu cơ bản (basic-needs)
- Học tập :
 - Thành tích học tập (academic-performance)
 - Tải học tập (study-load)
 - Mối quan hệ giáo viên – học sinh (teacher-student-relationship)
 - Quan ngại nghề nghiệp tương lai (future-career-concerns).

- Xã hội:
 - Hỗ trợ xã hội (social-support)
 - Áp lực từ bạn bè (peer-pressure)
 - Hoạt động ngoại khóa (extracurricular-activities)
 - Bắt nạt (bullying).
- stress-level (mức độ căng thẳng): là biến cuối cùng cho phép phân tích mối liên hệ giữa các yếu tố nói trên lên trạng thái căng thẳng của sinh viên.

Tiền xử lý dữ liệu

```
# Cài đặt và nạp thư viện
if (!require(MVN)) install.packages("MVN")
```

```
## Loading required package: MVN
```

```
library(MVN)
```

Đọc dữ liệu

```
data.set.1 <- read.csv('C:/Users/trand/OneDrive/Máy tính/Bao Cao Cuoi Ky/PROJECT/DATA/StressLevelDataset.csv')
# xem 6 dòng đầu tiên
head(data.set.1)
```

```
##   anxiety_level self_esteem mental_health_history depression headache
## 1             14          20                    0           11        2
## 2             15           8                    1           15        5
## 3             12          18                    1           14        2
## 4             16          12                    1           15        4
## 5             16          28                    0            7        2
## 6             20          13                    1           21        3
##   blood_pressure sleep_quality breathing_problem noise_level living_conditions
## 1                1              2                4            2                3
## 2                3              1                4            3                1
## 3                1              2                2            2                2
## 4                3              1                3            4                2
## 5                3              5                1            3                2
## 6                3              1                4            3                2
##   safety basic_needs academic_performance study_load
## 1      3          2              3            2
## 2      2          2              1            4
## 3      3          2              2            3
## 4      2          2              2            4
## 5      4          3              4            3
## 6      2          1              2            5
##   teacher_student_relationship future_career_concerns social_support
## 1                          3                      3            2
## 2                          1                      5            1
## 3                          3                      2            2
## 4                          1                      4            1
## 5                          1                      2            1
## 6                          2                      5            1
##   peer_pressure extracurricular_activities bullying stress_level
```

## 1	3	3	2	1
## 2	4	5	5	2
## 3	3	2	2	1
## 4	4	4	5	2
## 5	5	0	5	1
## 6	4	4	5	2

Thông tin bộ dữ liệu

Xem một số thông tin dữ liệu: - cấu trúc bộ dữ liệu: Quan trắc, Biến, Đặc điểm của biến. - Tóm tắt mô tả bộ dữ liệu.

```
cat("cấu trúc dữ liệu \n")
```

```
## cấu trúc dữ liệu
```

```
str(data.set.1)
```

```
## 'data.frame': 1100 obs. of 21 variables:
## $ anxiety_level : int 14 15 12 16 16 20 4 17 13 6 ...
## $ self_esteem : int 20 8 18 12 28 13 26 3 22 8 ...
## $ mental_health_history : int 0 1 1 1 0 1 0 1 1 0 ...
## $ depression : int 11 15 14 15 7 21 6 22 12 27 ...
## $ headache : int 2 5 2 4 2 3 1 4 3 4 ...
## $ blood_pressure : int 1 3 1 3 3 3 2 3 1 3 ...
## $ sleep_quality : int 2 1 2 1 5 1 4 1 2 1 ...
## $ breathing_problem : int 4 4 2 3 1 4 1 5 4 2 ...
## $ noise_level : int 2 3 2 4 3 3 1 3 3 0 ...
## $ living_conditions : int 3 1 2 2 2 2 4 1 3 5 ...
## $ safety : int 3 2 3 2 4 2 4 1 3 2 ...
## $ basic_needs : int 2 2 2 2 3 1 4 1 3 2 ...
## $ academic_performance : int 3 1 2 2 4 2 5 1 3 2 ...
## $ study_load : int 2 4 3 4 3 5 1 3 3 2 ...
## $ teacher_student_relationship: int 3 1 3 1 1 2 4 2 2 1 ...
## $ future_career_concerns : int 3 5 2 4 2 5 1 4 3 5 ...
## $ social_support : int 2 1 2 1 1 1 3 1 3 1 ...
## $ peer_pressure : int 3 4 3 4 5 4 2 4 3 5 ...
## $ extracurricular_activities : int 3 5 2 4 0 4 2 4 2 3 ...
## $ bullying : int 2 5 2 5 5 5 1 5 2 4 ...
## $ stress_level : int 1 2 1 2 1 2 0 2 1 1 ...
```

Tóm tắt mô tả bộ dữ liệu

```
summary(data.set.1)
```

```
## anxiety_level self_esteem mental_health_history depression
## Min. : 0.00 Min. : 0.00 Min. :0.0000 Min. : 0.00
## 1st Qu.: 6.00 1st Qu.:11.00 1st Qu.:0.0000 1st Qu.: 6.00
## Median :11.00 Median :19.00 Median :0.0000 Median :12.00
## Mean :11.06 Mean :17.78 Mean :0.4927 Mean :12.56
## 3rd Qu.:16.00 3rd Qu.:26.00 3rd Qu.:1.0000 3rd Qu.:19.00
## Max. :21.00 Max. :30.00 Max. :1.0000 Max. :27.00
## headache blood_pressure sleep_quality breathing_problem
## Min. :0.000 Min. :1.000 Min. :0.00 Min. :0.000
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:1.00 1st Qu.:2.000
## Median :3.000 Median :2.000 Median :2.50 Median :3.000
## Mean :2.508 Mean :2.182 Mean :2.66 Mean :2.754
```

```
## 3rd Qu.:3.000 3rd Qu.:3.000 3rd Qu.:4.00 3rd Qu.:4.000
## Max. :5.000 Max. :3.000 Max. :5.00 Max. :5.000
## noise_level living_conditions safety basic_needs
## Min. :0.000 Min. :0.000 Min. :0.000 Min. :0.000
## 1st Qu.:2.000 1st Qu.:2.000 1st Qu.:2.000 1st Qu.:2.000
## Median :3.000 Median :2.000 Median :2.000 Median :3.000
## Mean :2.649 Mean :2.518 Mean :2.737 Mean :2.773
## 3rd Qu.:3.000 3rd Qu.:3.000 3rd Qu.:4.000 3rd Qu.:4.000
## Max. :5.000 Max. :5.000 Max. :5.000 Max. :5.000
## academic_performance study_load teacher_student_relationship
## Min. :0.000 Min. :0.000 Min. :0.000
## 1st Qu.:2.000 1st Qu.:2.000 1st Qu.:2.000
## Median :2.000 Median :2.000 Median :2.000
## Mean :2.773 Mean :2.622 Mean :2.648
## 3rd Qu.:4.000 3rd Qu.:3.000 3rd Qu.:4.000
## Max. :5.000 Max. :5.000 Max. :5.000
## future_career_concerns social_support peer_pressure
## Min. :0.000 Min. :0.000 Min. :0.000
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:2.000
## Median :2.000 Median :2.000 Median :2.000
## Mean :2.649 Mean :1.882 Mean :2.735
## 3rd Qu.:4.000 3rd Qu.:3.000 3rd Qu.:4.000
## Max. :5.000 Max. :3.000 Max. :5.000
## extracurricular_activities bullying stress_level
## Min. :0.000 Min. :0.000 Min. :0.0000
## 1st Qu.:2.000 1st Qu.:1.000 1st Qu.:0.0000
## Median :2.500 Median :3.000 Median :1.0000
## Mean :2.767 Mean :2.617 Mean :0.9964
## 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.:2.0000
## Max. :5.000 Max. :5.000 Max. :2.0000
```

Nhận xét: Nhìn tổng quan ta thấy có 21 biến và 1100 quan trắc. Ở đây ta thấy rằng cả 21 biến ở đây đều là biến số.

Xem giá trị hiện diện ở các biến

```
lapply(data.set.1, function(x) sort(unique(x)))
```

```
## $anxiety_level
## [1] 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
##
## $self_esteem
## [1] 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
## [26] 25 26 27 28 29 30
##
## $mental_health_history
## [1] 0 1
##
## $depression
## [1] 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
## [26] 25 26 27
##
## $headache
## [1] 0 1 2 3 4 5
```

```

##
## $blood_pressure
## [1] 1 2 3
##
## $sleep_quality
## [1] 0 1 2 3 4 5
##
## $breathing_problem
## [1] 0 1 2 3 4 5
##
## $noise_level
## [1] 0 1 2 3 4 5
##
## $living_conditions
## [1] 0 1 2 3 4 5
##
## $safety
## [1] 0 1 2 3 4 5
##
## $basic_needs
## [1] 0 1 2 3 4 5
##
## $academic_performance
## [1] 0 1 2 3 4 5
##
## $study_load
## [1] 0 1 2 3 4 5
##
## $teacher_student_relationship
## [1] 0 1 2 3 4 5
##
## $future_career_concerns
## [1] 0 1 2 3 4 5
##
## $social_support
## [1] 0 1 2 3
##
## $peer_pressure
## [1] 0 1 2 3 4 5
##
## $extracurricular_activities
## [1] 0 1 2 3 4 5
##
## $bullying
## [1] 0 1 2 3 4 5
##
## $stress_level
## [1] 0 1 2

```

Nhận xét: Trong đó, hầu như các biến là biến thang đo chỉ mức độ và đồng thời một số biến có thang đo thấp như là tiền sử sức khỏe tâm thần 0 – 1, cao huyết áp 1 – 3, ... quá ít để đánh có thể giả định chuẩn đa biến và đồng nhất phương sai giữa các nhóm được (2 điều kiện trong phân tích MANOVA). Nên vì vậy trong bài có thể phân tích PCA và FA (đối với FA vì ta đang phân tích trên mẫu lớn nên có thể phân tích được)

Kết luận: Với bộ dữ liệu ta sẽ thực hiện phân tích với 20 biến đầu tiên và biến cuối cùng stress_level sẽ là

biến phân loại

```
# chuyển biến stress_level thành biến phân loại:
data.set.1$stress_level <- factor(data.set.1$stress_level)
```

Tiền xử lý dữ liệu

Kiểm tra dữ liệu thiếu

```
missing_values <- colSums(is.na(data.set.1))
print("số dữ liệu thiếu trong mỗi biến: ")
```

```
## [1] "số dữ liệu thiếu trong mỗi biến: "
```

```
print(missing_values)
```

```
##          anxiety_level          self_esteem
##                0                0
##    mental_health_history          depression
##                0                0
##          headache          blood_pressure
##                0                0
##    sleep_quality          breathing_problem
##                0                0
##    noise_level          living_conditions
##                0                0
##          safety          basic_needs
##                0                0
##    academic_performance          study_load
##                0                0
## teacher_student_relationship    future_career_concerns
##                0                0
##          social_support          peer_pressure
##                0                0
##    extracurricular_activities          bullying
##                0                0
##          stress_level
##                0
```

Nhận xét Bộ dữ liệu không có giá trị thiếu, đảm bảo tính đầy đủ cho phân tích.

Ta thực hiện lấy 20 biến đầu tiên để phân tích (biến yếu tố - major) - kiểm tra tính chuẩn

```
data.major <- as.data.frame(data.set.1[,1:20])
```

```
names <- colnames(data.set.1)[1:20]
sapply(data.major[names], class) # Kiểm tra kiểu dữ liệu
```

```
##          anxiety_level          self_esteem
##          "integer"          "integer"
##    mental_health_history          depression
##          "integer"          "integer"
##          headache          blood_pressure
##          "integer"          "integer"
##    sleep_quality          breathing_problem
##          "integer"          "integer"
##    noise_level          living_conditions
##          "integer"          "integer"
```

```
##             safety                basic_needs
##             "integer"              "integer"
## academic_performance                study_load
##             "integer"              "integer"
## teacher_student_relationship        future_career_concerns
##             "integer"              "integer"
##             social_support          peer_pressure
##             "integer"              "integer"
## extracurricular_activities          bullying
##             "integer"              "integer"
```

Nhận xét: Các biến số đều là biến số interger (là biến số)

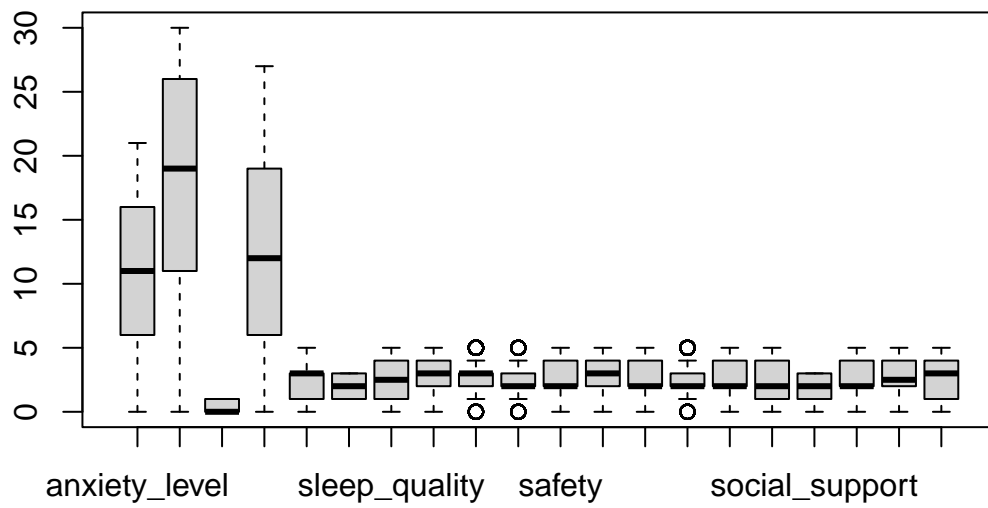
```
# 7. Kiểm tra giá trị bất thường (outliers)
# Ví dụ: Kiểm tra giá trị âm hoặc bất thường
outliers_check <- function(column) {
  boxplot.stats(column)$out
}
for (colname in names(data.major)) {
  cat("Outliers in", colname, ":", outliers_check(data.major[[colname]]), "\n")
}
```

```
## Outliers in anxiety_level :
## Outliers in self_esteem :
## Outliers in mental_health_history :
## Outliers in depression :
## Outliers in headache :
## Outliers in blood_pressure :
## Outliers in sleep_quality :
## Outliers in breathing_problem :
## Outliers in noise_level : 0 5 5 5 5 5 5 5 5 5 5 5 5 5 0 0 0 5 5 5 5 5 5 5 5 5 0 0 5 5 5 0 5 5 5
## Outliers in living_conditions : 5 5 5 5 0 5 0 0 0 5 0 0 5 5 5 0 5 5 0 0 0 5 5 0 0 5 5 0 0 5 0 5 0 5
## Outliers in safety :
## Outliers in basic_needs :
## Outliers in academic_performance :
## Outliers in study_load : 5 5 5 0 0 0 5 5 5 5 5 5 5 0 5 5 5 5 5 5 5 5 0 5 5 0 5 0 5 5 0 0 5 5 5 0
## Outliers in teacher_student_relationship :
## Outliers in future_career_concerns :
## Outliers in social_support :
## Outliers in peer_pressure :
## Outliers in extracurricular_activities :
## Outliers in bullying :
```

Nhận xét: ta nhận thấy chỉ có biến noise_level, living_conditions, study_load, là có outlier, ta sẽ xem xét thử xem bộ dữ liệu có outlier đa biến hay không.

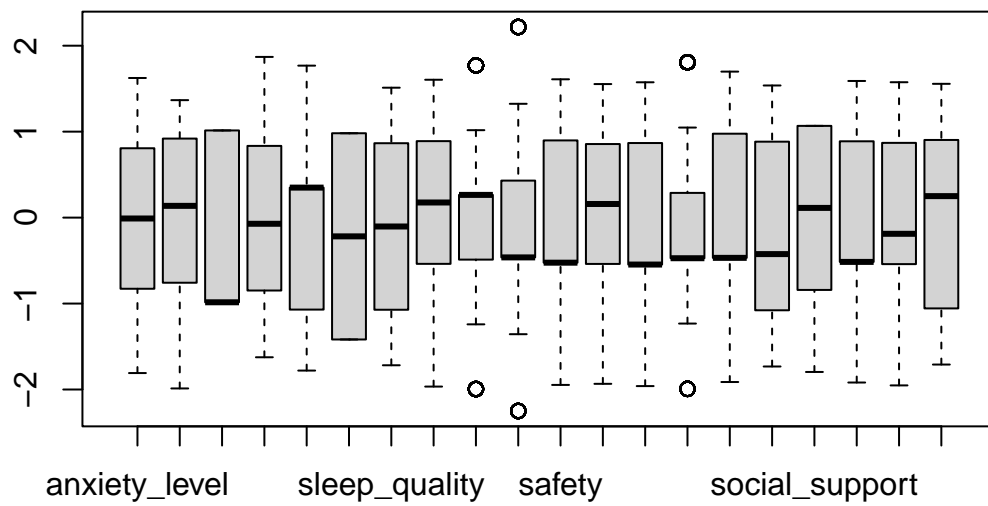
Xem qua chút về đồ thị boxplot của 20 biến đầu tiên:

```
boxplot(data.major)
```



tiến hành chuẩn hóa bộ dữ liệu

```
data.major.sc <- as.data.frame(scale(data.major, center = T, scale = T))
boxplot(data.major.sc)
```



xét tính chuẩn đa biến của bộ dữ liệu

Loại bỏ các dữ liệu là outlier đa biến

```
x_mean <- colMeans(data.major.sc)
S <- cov(data.major.sc)
d2 <- mahalanobis(data.major.sc, x_mean, S)

# Xác định ngưỡng loại bỏ (975% phân phối Chi-squared)
cutoff <- qchisq(0.975, df = ncol(data.major.sc))

# số giá trị ngoại lai trong bộ dữ liệu:
cat("số giá trị ngoại lai đa biến trong bộ dữ liệu:")
```

số giá trị ngoại lai đa biến trong bộ dữ liệu:

```
(length(which(d2 >= cutoff)))
```

[1] 200

```
# Loại bỏ outliers
non_outlier_index <- which(d2 < cutoff)
data.set.1.clean <- data.set.1[non_outlier_index, ]
cat("Outlier đã được loại bỏ")
```

Outlier đã được loại bỏ

Kiểm tra số quan trắc của bộ dữ liệu sau khi lọc outlier

```
dim(data.set.1.clean)
```

[1] 900 21

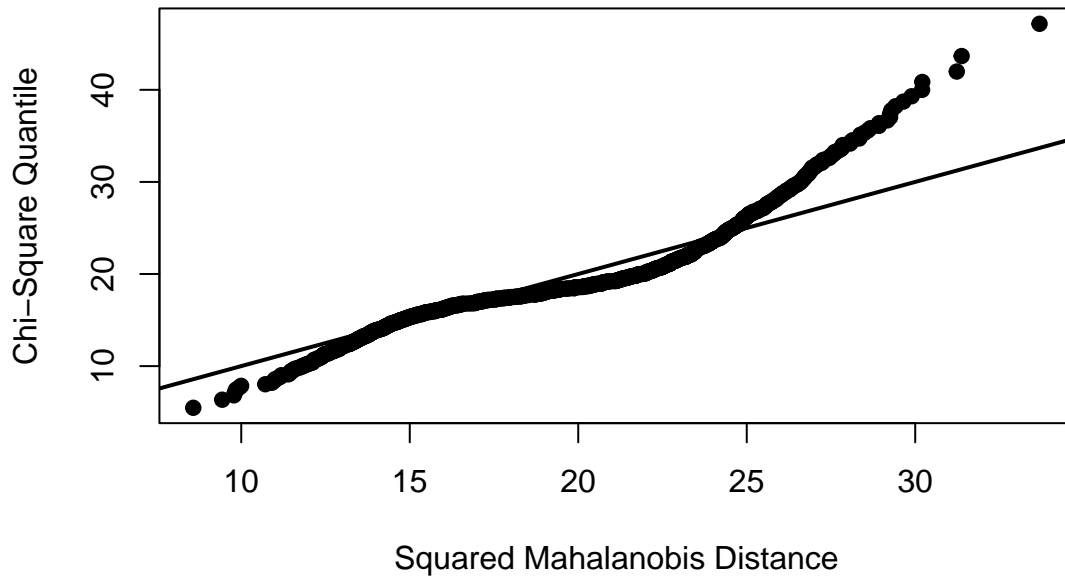
Nhận xét: bộ dữ liệu còn lại 900 quan trắc và 21 biến

kiểm tra tính chuẩn đa biến

kiểm tra lại tính chuẩn đa biến của bộ dữ liệu sau khi lọc outlier (lấy 20 biến đầu tiên ra chứa lại biến phân loại stress_level)

```
data.major <- data.set.1.clean[,1:20]
data.major.sc <- as.data.frame(scale(data.major, scale = T, center = T))
result <- mvn(data.major.sc, mvnTest = "mardia",multivariatePlot = "qq")
```

Chi-Square Q-Q Plot



```
print(result$multivariateNormality)
```

##	Test	Statistic	p value	Result
## 1	Mardia Skewness	4781.78190699248	0	NO
## 2	Mardia Kurtosis	-7.62603770472107	2.42028619368284e-14	NO
## 3	MVN	<NA>	<NA>	NO

Nhận xét: bộ dữ liệu không có tính chuẩn đa biến kể cả khi nhìn vào thống kê lần đồ thị qq. ta sẽ xem thử các biến có tính chuẩn đơn biến hay không.

```
print(result$univariateNormality)
```

##	Test	Variable	Statistic	p value	Normality
## 1	Anderson-Darling	anxiety_level	7.6989	<0.001	NO
## 2	Anderson-Darling	self_esteem	19.6091	<0.001	NO
## 3	Anderson-Darling	mental_health_history	161.5289	<0.001	NO
## 4	Anderson-Darling	depression	8.5589	<0.001	NO
## 5	Anderson-Darling	headache	41.8001	<0.001	NO
## 6	Anderson-Darling	blood_pressure	73.5567	<0.001	NO
## 7	Anderson-Darling	sleep_quality	46.7528	<0.001	NO
## 8	Anderson-Darling	breathing_problem	42.9065	<0.001	NO
## 9	Anderson-Darling	noise_level	33.7638	<0.001	NO
## 10	Anderson-Darling	living_conditions	42.1861	<0.001	NO
## 11	Anderson-Darling	safety	40.5067	<0.001	NO
## 12	Anderson-Darling	basic_needs	39.3523	<0.001	NO
## 13	Anderson-Darling	academic_performance	40.1607	<0.001	NO
## 14	Anderson-Darling	study_load	34.5314	<0.001	NO
## 15	Anderson-Darling	teacher_student_relationship	38.4714	<0.001	NO
## 16	Anderson-Darling	future_career_concerns	47.1548	<0.001	NO
## 17	Anderson-Darling	social_support	112.4160	<0.001	NO
## 18	Anderson-Darling	peer_pressure	41.5523	<0.001	NO

```
## 19 Anderson-Darling  extracurricular_activities  38.7093 <0.001    NO
## 20 Anderson-Darling      bullying                44.7343 <0.001    NO
```

Nhận xét: dựa vào kiểm định ta nhận thấy không có biến trong bộ 20 biến đầu tiên của bộ dữ liệu có tính chuẩn đơn biến.

Bộ dữ liệu hoàn chỉnh sau khi đã làm sạch - chưa chuẩn hóa: (có chuyển đổi biến stress_level thành biến factor)

```
data.set.1 <- data.set.1.clean
```

Phương pháp PCA

```
if (!require(ggplot2)) install.packages("ggplot2")
```

```
## Loading required package: ggplot2
```

```
if (!require(factoextra)) install.packages("factoextra")
```

```
## Loading required package: factoextra
```

```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```
if (!require(ggfortify)) install.packages("ggfortify")
```

```
## Loading required package: ggfortify
```

```
library(factoextra)
```

```
library(ggplot2)
```

```
library(ggfortify)
```

Phân định số lượng nhân tố giữ lại

Lấy 20 biến đầu tiên để phân tích:

```
data.major <- data.set.1[,1:20]
```

Thực hiện pca trên bộ dữ liệu theo ma trận hệ số tương quan (bộ dữ liệu đã được chuẩn hóa):

```
pc.dat <- princomp(data.major, cor = T)
```

PHƯƠNG PHÁP TỶ LỆ PHƯƠNG SAI TOÀN PHẦN

```
summary(pc.dat)
```

```
## Importance of components:
```

```
##              Comp.1      Comp.2      Comp.3      Comp.4      Comp.5
## Standard deviation  3.978417  1.05911964  0.60932484  0.58420674  0.52917619
## Proportion of Variance 0.791390  0.05608672  0.01856384  0.01706488  0.01400137
## Cumulative Proportion 0.791390  0.84747674  0.86604058  0.88310545  0.89710682
##              Comp.6      Comp.7      Comp.8      Comp.9
## Standard deviation  0.51438402  0.47741909  0.416182223  0.397319720
## Proportion of Variance 0.01322955  0.01139645  0.008660382  0.007893148
## Cumulative Proportion 0.91033637  0.92173282  0.930393200  0.938286348
##              Comp.10     Comp.11     Comp.12     Comp.13
## Standard deviation  0.389917516  0.382534583  0.370793419  0.366817359
## Proportion of Variance 0.007601783  0.007316635  0.006874388  0.006727749
## Cumulative Proportion 0.945888132  0.953204767  0.960079155  0.966806904
##              Comp.14     Comp.15     Comp.16     Comp.17
## Standard deviation  0.356934661  0.341278403  0.336573427  0.299853275
```

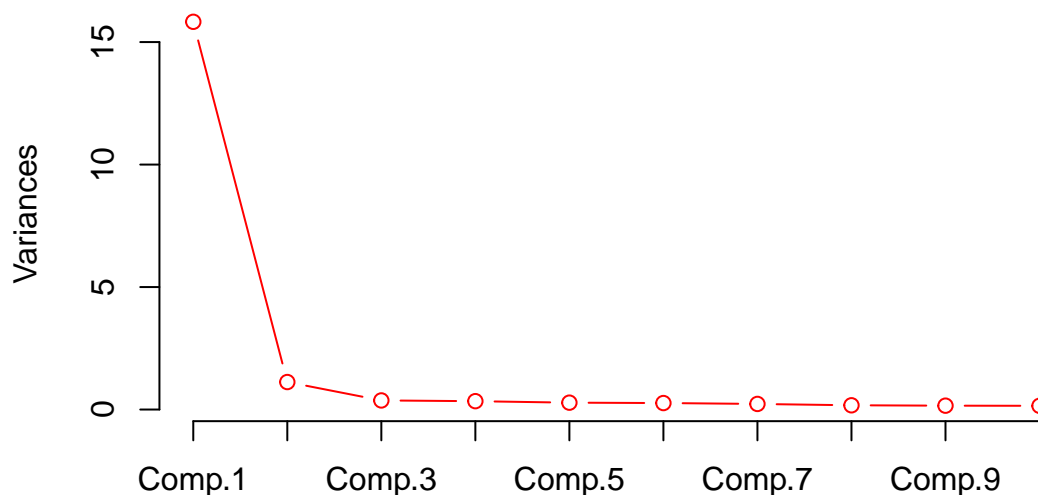
```
## Proportion of Variance 0.006370118 0.005823547 0.005664084 0.004495599
## Cumulative Proportion 0.973177021 0.979000569 0.984664653 0.989160252
##                               Comp.18      Comp.19      Comp.20
## Standard deviation      0.288318186 0.265764591 0.251071244
## Proportion of Variance 0.004156369 0.003531541 0.003151838
## Cumulative Proportion 0.993316621 0.996848162 1.000000000
```

Nhận xét: dựa trên tỷ lệ phương sai toàn phần ta dựa vào cột Cumulative Proportion thì ta nhận thấy rằng 2 nhân tố đầu tiên đã chiếm đến 84.7% thông tin của bộ dữ liệu, 3 nhân tố đầu tiên là 86.6%, sự chênh lệch giữa 3 nhân tố và 2 nhân tố là 1.85%, sự chênh lệch này là không quá lớn nên có lẽ ta sẽ giữ lại 2 nhân tố đầu tiên. Nhưng để đánh giá rõ ràng vì kĩ hơn ta sẽ xem xét thêm phương pháp dựa trên trị riêng và phương pháp khử tay.

PHƯƠNG PHÁP DỰA TRÊN TRỊ RIÊNG và đồ thị khử tay

```
plot(pc.dat, type = "line", col = "red", main = "ĐỒ THỊ KHỬ TAY DỰA TRÊN THÀNH PHẦN CHÍNH")
```

Đ. TH. KH.Y TAY D.A TRÊN THÀNH PH.N CHÍNH



```
(pc.dat$sdev^2)
```

```
##      Comp.1      Comp.2      Comp.3      Comp.4      Comp.5      Comp.6
## 15.82780033  1.12173442  0.37127676  0.34129751  0.28002744  0.26459092
##      Comp.7      Comp.8      Comp.9      Comp.10     Comp.11     Comp.12
##  0.22792899  0.17320764  0.15786296  0.15203567  0.14633271  0.13748776
##      Comp.13     Comp.14     Comp.15     Comp.16     Comp.17     Comp.18
##  0.13455498  0.12740235  0.11647095  0.11328167  0.08991199  0.08312738
##      Comp.19     Comp.20
##  0.07063082  0.06303677
```

Nhận xét: Dựa trên phương pháp này ta sẽ nhận những trị riêng lớn hơn 1, ở đây thì ta nhận thấy rằng có 2 thành phần đầu tiên là có trị riêng lớn hơn 1 và đồng thời dựa vào biểu đồ khử tay thì đồ thị bị gấp khúc ở thành phần thứ 2, từ 2 trở đi thì sự gấp khúc không đáng kể. Tổng kết dựa vào 2 phương pháp ta càng

khẳng định ta giữ lại 2 thành phần đầu tiên.

Hệ số tải của 3 thành phần chính đầu tiên

```
round(pc.dat$loadings[,1:3],2)
```

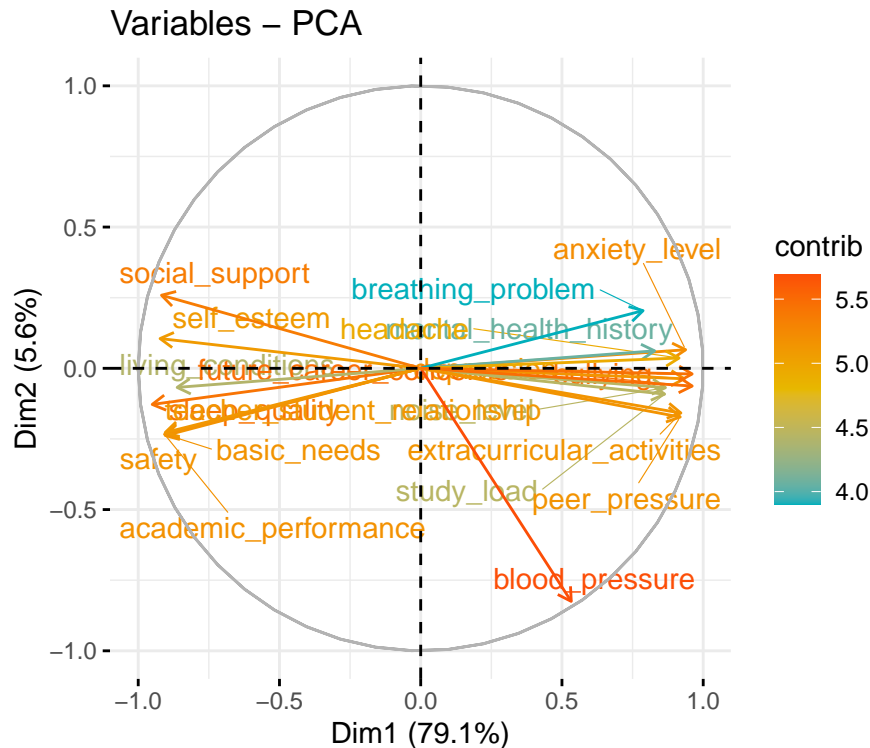
##	Comp.1	Comp.2	Comp.3
## anxiety_level	0.24	0.06	0.05
## self_esteem	-0.23	0.10	0.01
## mental_health_history	0.21	0.06	0.23
## depression	0.23	-0.03	0.00
## headache	0.23	0.03	0.11
## blood_pressure	0.13	-0.78	-0.11
## sleep_quality	-0.24	-0.12	-0.09
## breathing_problem	0.20	0.19	-0.93
## noise_level	0.22	-0.06	0.06
## living_conditions	-0.22	-0.06	0.06
## safety	-0.23	-0.22	-0.05
## basic_needs	-0.23	-0.22	-0.08
## academic_performance	-0.23	-0.22	-0.11
## study_load	0.22	-0.09	0.01
## teacher_student_relationship	-0.23	-0.21	-0.10
## future_career_concerns	0.24	-0.06	0.03
## social_support	-0.23	0.24	-0.01
## peer_pressure	0.23	-0.16	-0.04
## extracurricular_activities	0.23	-0.15	0.04
## bullying	0.24	-0.02	0.03

Nhận xét:

- Thành phần chính 1 (Comp.1): ta nhận thấy dường như các hệ số các biến tương đương nhau về mặt độ lớn, điều đó cho thấy ở thành phần này biểu diễn toàn bộ các ý nghĩa của 20 biến yếu tố của bộ dữ liệu, hệ số dương biểu diễn ý nghĩa tương quan thuận còn âm biểu diễn ý nghĩa tương quan nghịch. Trong đó có 1 biến mà ta thấy rõ chênh lệch so với các biến còn lại là biến blood_pressure với hệ số 0.13, dường như điều này cho thấy ở biến cao huyết áp ở thành phần này biểu diễn không rõ ràng lắm.
- Thành phần chính 2 (Comp.2): ta nhận thấy dường như biến blood_pressure là có hệ số lớn nhất với lên đến 0.78 về mặt độ lớn, điều này có nghĩa ở thành phần này dường như đang nói lên ý nghĩa tiềm ẩn chính là cao huyết áp có liên quan mạnh đến stress_level.
- Thành phần chính 3 (comp.3): tương tự như ở thành phần chính thứ 2 thì ở đây là biến breathing_problem lên đến 0.98, điều này cho thấy có vẻ như vấn đề về hô hấp cũng là một nhân tố ẩn trong đây

Những điều ta phân tích trên đều thể hiện rõ ràng qua đồ thị hệ số tải:

```
fviz_pca_var(pc.dat,  
             col.var = "contrib", # Color by contributions to the PC  
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),  
             repel = TRUE        # Avoid text overlapping  
)
```



Kết luận: ở phân tích thành phần chính ta sẽ quan tâm chủ yếu ở 2 thành phần đầu tiên như ta đã đề cập, còn ý nghĩa về 2 biến ẩn ta sẽ phân tích ở FA sau.

Phân tích dựa trên biểu đồ score

Xem qua biểu đồ score chính là dựa trên bộ dữ liệu mới với các điểm dữ liệu mới dựa trên các thành phần chính mà ta đã phân tích:

vì ta giữ lại 2 thành phần chính đầu tiên nên ta sẽ lấy score của 2 thành phần chính đầu tiên:

```
score <- as.data.frame(pc.dat$scores[,1:2])
```

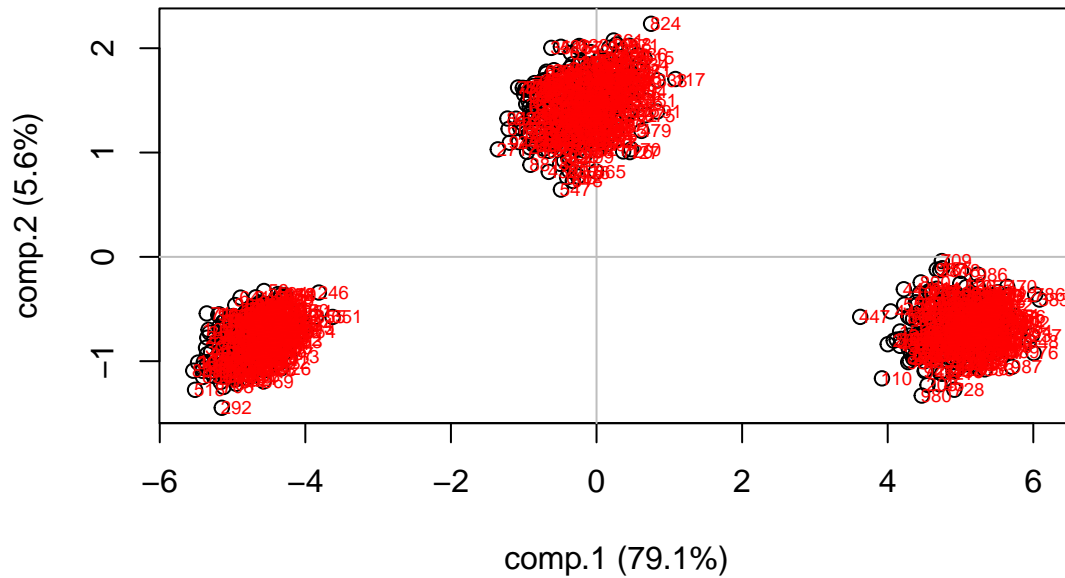
Ta thêm biến phân loại stress_level vào bộ dữ liệu mới này:

```
score$stress_level <- data.set.1$stress_level
```

Ta sẽ xem qua biểu đồ phân tán của 2 thành phần chính đầu tiên:

```
plot(x = score[,1], y = score[,2], xlab = "comp.1 (79.1%)", ylab = "comp.2 (5.6%)", main = "Đồ thị Score")
abline(h = 0, v = 0, col = "gray")
# vẽ thêm số quan trắc của score
text(score[, 1] + 0.2, score[, 2], label = rownames(score), col="red", cex=0.6)
```

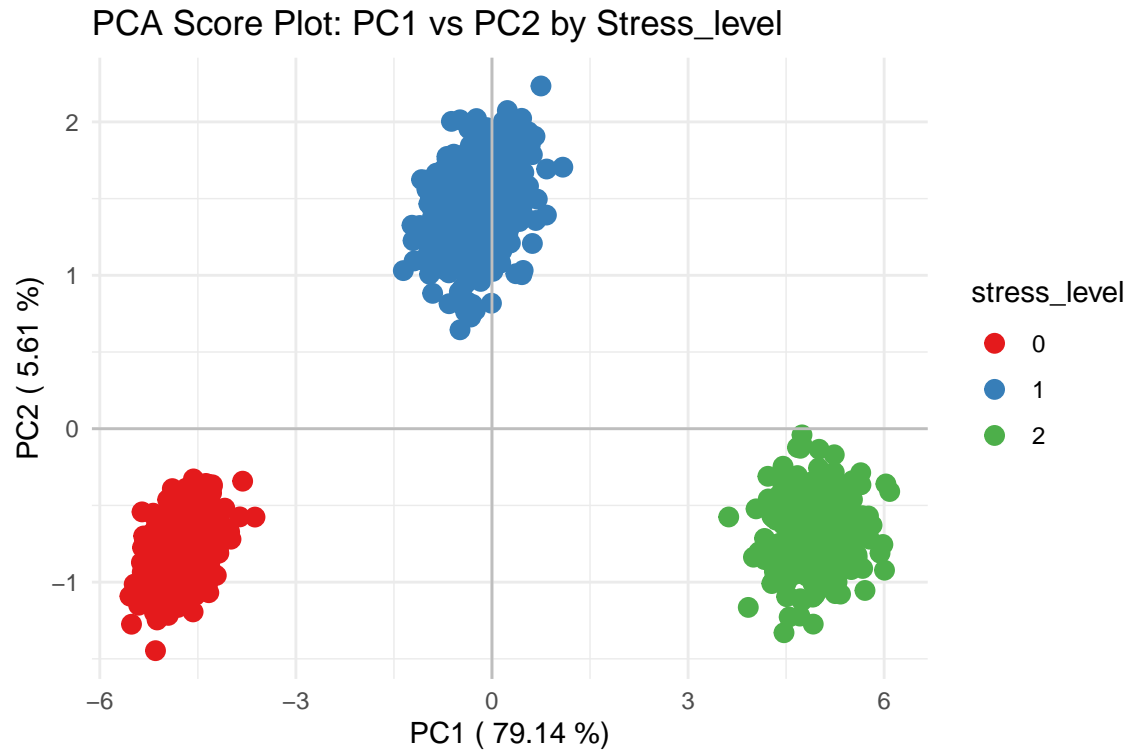
Đ. th. Scores



Nhận xét: xem qua biểu đồ score ta nhận thấy rất rõ ràng 3 điểm vị trí tách biệt với nhau, dựa trên những phán đoán ban đầu có vẻ như 3 vị trí này sẽ là 3 vị trí phân loại của 3 mức độ stress mà được thể hiện trong biến stress_level

Vẽ biểu đồ score theo biến phân loại stress pc1 và pc2

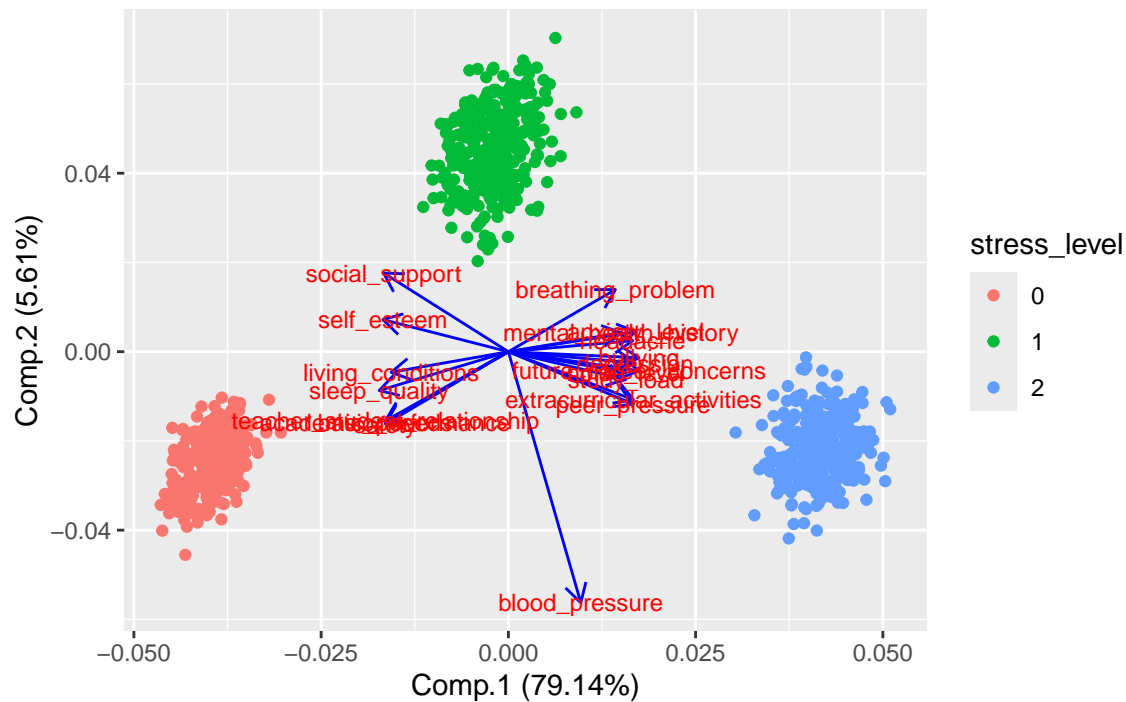
```
ggplot(score, aes(x = Comp.1, y = Comp.2, color = stress_level)) +
  geom_point(size = 3) +
  labs(title = "PCA Score Plot: PC1 vs PC2 by Stress_level",
        x = paste("PC1 (", round((pc.dat$sdev[1])^2/sum((pc.dat$sdev)^2)*100, 2), "%)"),
        y = paste("PC2 (", round((pc.dat$sdev[2])^2/sum((pc.dat$sdev)^2)*100, 2), "%)")) +
  geom_vline(xintercept = 0, color = "gray", linetype = "solid") + # Trục y=0 +
  geom_hline(yintercept = 0, color = "gray", linetype = "solid") + # Trục x=0 +
  theme_minimal() +
  scale_color_brewer(palette = "Set1")
```



Nhận xét: đúng như những gì ta dự đoán, dựa vào đồ thị score theo biến phân loại stress_level ta nhận định được rõ ràng mức độ stress đã được thể hiện qua 3 vùng, ta sẽ kết hợp cùng với đồ thị loadings để đánh giá những yếu tố nào là yếu tố ràng khiến cho sự phân loại mức độ stress thể hiện rõ như vậy.

```
autoplot(pc.dat, loadings = TRUE, loadings.colour = 'blue', loadings.label = TRUE, loadings.label.size = 10,
         colour = "stress_level", data = data.set.1, main="Biểu đồ score cho PC1 và PC2")
```


Biểu đồ score cho PC1 và PC2



Nhận xét: dựa vào đồ thị trên ta nhận định rõ ràng được việc những quan trắc mà có mức độ stress là 0 được thể hiện ở các biến có hệ số tải âm trong thành phần 1, vì vậy ta sẽ tiến hành trích xuất tên các biến đó:

MỨC ĐỘ STRESS = 0

```
pc.loading <- pc.dat$loadings[,1:2]
cat("Tên các yếu tố tác động để nhận định người có mức độ stress là 0\n")
```

```
## Tên các yếu tố tác động để nhận định người có mức độ stress là 0
```

```
cat("Những người có các yếu tố sau đây lớn:\n")
```

```
## Những người có các yếu tố sau đây lớn:
```

```
names(which(pc.loading[,1] < 0))
```

```
## [1] "self_esteem"           "sleep_quality"
## [3] "living_conditions"     "safety"
## [5] "basic_needs"           "academic_performance"
## [7] "teacher_student_relationship" "social_support"
```

```
cat("Những người có các yếu tố sau đây thấp:\n")
```

```
## Những người có các yếu tố sau đây thấp:
```

```
names(which(pc.loading[,1] > 0))
```

```
## [1] "anxiety_level"         "mental_health_history"
## [3] "depression"            "headache"
## [5] "blood_pressure"        "breathing_problem"
## [7] "noise_level"           "study_load"
## [9] "future_career_concerns" "peer_pressure"
## [11] "extracurricular_activities" "bullying"
```

Nhận xét: ta nhận thấy được những người mà có lòng tự trọng (self_esteem), sleep_quality (chất lượng giấc ngủ), living_conditions (điều kiện sống), điều kiện sống (safety), nhu cầu cơ bản(basic_needs), thành tích học tập (academic_performance), quan hệ giữa giáo viên và học sinh (teacher_student_relationship), hỗ trợ xã hội (social_support) cao thì là những người có mức độ stress là 0. Đồng thời mức độ lo lắng(anxiety_level), tiền sử sức khỏe tâm thần (mental_health_history), trầm cảm(depression), đau đầu(headache), huyết áp (blood_pressure), vấn đề hô hấp (breathing_problem), mức độ ồn(noise_level), tải học tập (study_load), quan ngại về nghề nghiệp tương lai(future_career_concerns), áp lực đồng trang lứa(peer_pressure), hoạt động ngoại khóa(extracurricular_activities), bắt nạt(bullying) thì các yếu tố này cũng giúp phân định một người có mức độ stress là 0

MỨC ĐỘ STRESS = 2

```
cat("Tên các yếu tố tác động để nhận định người có mức độ stress là 2\n")
```

```
## Tên các yếu tố tác động để nhận định người có mức độ stress là 2
```

```
cat("Những người có các yếu tố sau đây lớn:\n")
```

```
## Những người có các yếu tố sau đây lớn:
```

```
names(which(pc.loading[,1] > 0 ))
```

```
## [1] "anxiety_level"          "mental_health_history"
## [3] "depression"             "headache"
## [5] "blood_pressure"         "breathing_problem"
## [7] "noise_level"            "study_load"
## [9] "future_career_concerns" "peer_pressure"
## [11] "extracurricular_activities" "bullying"
```

```
cat("Những người có các yếu tố sau đây thấp:\n")
```

```
## Những người có các yếu tố sau đây thấp:
```

```
names(which(pc.loading[,1] < 0))
```

```
## [1] "self_esteem"           "sleep_quality"
## [3] "living_conditions"     "safety"
## [5] "basic_needs"           "academic_performance"
## [7] "teacher_student_relationship" "social_support"
```

Nhận xét: ta nhận thấy được những người mà có mức độ lo lắng(anxiety_level), tiền sử sức khỏe tâm thần (mental_health_history), trầm cảm(depression), đau đầu(headache), huyết áp (blood_pressure), vấn đề hô hấp (breathing_problem), mức độ ồn(noise_level), tải học tập (study_load), quan ngại về nghề nghiệp tương lai(future_career_concerns), áp lực đồng trang lứa(peer_pressure), hoạt động ngoại khóa(extracurricular_activities), bắt nạt(bullying) cao thì các yếu tố này cũng giúp phân định một người có mức độ stress là 0. đồng thời lòng tự trọng (self_esteem), sleep_quality (chất lượng giấc ngủ), living_conditions (điều kiện sống), điều kiện sống (safety), nhu cầu cơ bản(basic_needs), thành tích học tập (academic_performance), quan hệ giữa giáo viên và học sinh (teacher_student_relationship), hỗ trợ xã hội (social_support) thấp thì là những người có mức độ stress là 0

Tổng kết: Phân tích PCA cho thấy chỉ cần 2 thành phần chính đầu tiên là đủ để biểu diễn phần lớn thông tin dữ liệu (84.7%) và giúp phân loại rõ ràng mức độ stress. Các thành phần này kết hợp thông tin từ nhiều biến ban đầu, giúp rút gọn dữ liệu và phát hiện các yếu tố tiềm ẩn ảnh hưởng mạnh đến sức khỏe tinh thần, như:

- Nhóm yếu tố tích cực (tự trọng, giấc ngủ, hỗ trợ xã hội).
- Nhóm yếu tố tiêu cực (lo âu, trầm cảm, bắt nạt, tải học tập, huyết áp)

Những yếu tố này sẽ đóng vai trò quan trọng trong các bước phân tích nhân tố (FA) tiếp theo.

Phương pháp FA

```
if (!require(psych)) install.packages("psych")
```

```
## Loading required package: psych
```

```
##
```

```
## Attaching package: 'psych'
```

```
## The following objects are masked from 'package:ggplot2':
```

```
##
```

```
##      %+%, alpha
```

```
library(psych)
```

FA với phương pháp Maximum Likelihood và quay varimax

```
fa.result <- fa(r = data.major, nfactors = 2, rotate = "varimax", fm = "ml", scores = "Bartlett")
```

Giá trị loadings

```
print(fa.result$loadings, cutoff = 0.3)
```

```
##
```

```
## Loadings:
```

##	ML2	ML1
## anxiety_level	0.867	0.354
## self_esteem	-0.787	-0.484
## mental_health_history	0.753	0.316
## depression	0.822	0.436
## headache	0.835	0.363
## blood_pressure		0.991
## sleep_quality	-0.912	-0.301
## breathing_problem	0.755	
## noise_level	0.747	0.424
## living_conditions	-0.790	-0.322
## safety	-0.907	
## basic_needs	-0.906	
## academic_performance	-0.909	
## study_load	0.741	0.429
## teacher_student_relationship	-0.902	
## future_career_concerns	0.840	0.470
## social_support	-0.719	-0.617
## peer_pressure	0.753	0.541
## extracurricular_activities	0.763	0.527
## bullying	0.858	0.435

```
##
```

```
##              ML2    ML1
```

```
## SS loadings  12.848 3.886
```

```
## Proportion Var 0.642 0.194
```

```
## Cumulative Var 0.642 0.837
```

- Nhận xét ban đầu:
 - Hai nhân tố giải thích được 83.7% phương sai, cho thấy mô hình FA với 2 nhân tố khá hiệu quả trong việc nắm bắt cấu trúc tiềm ẩn của dữ liệu.
 - ML2 đóng vai trò chính (64.2%), trong khi ML1 bổ sung thêm 19.4%.
- Nhân tố 1 (ML1)
 - Biến có tải trọng cao:

- * blood_pressure (0.991): Tải trọng gần 1, cho thấy nhân tố này gần như hoàn toàn đại diện cho huyết áp.
- * Các biến khác có tải trọng trung bình: anxiety_level (0.354), depression (0.436), headache (0.363), noise_level (0.424), study_load (0.429), future_career_concerns (0.470), peer_pressure (0.541), extracurricular_activities (0.527), bullying (0.435).
- Biến có tải trọng thấp hoặc âm:
 - * self_esteem (-0.484), sleep_quality (-0.301), living_conditions (-0.322), social_support (-0.617) có tải trọng âm, nhưng không mạnh. -Diễn giải:
 - * ML1 chủ yếu được định hình bởi blood_pressure, cho thấy nhân tố này đại diện cho yếu tố sinh lý (đặc biệt là huyết áp, một chỉ số thường tăng khi căng thẳng). Các biến như anxiety_level, depression, peer_pressure, bullying có tải trọng trung bình, gợi ý rằng nhân tố này cũng phản ánh một phần căng thẳng tâm lý và xã hội, nhưng bị lấn át bởi blood_pressure.
 - * Tải trọng âm của self_esteem và social_support cho thấy mối quan hệ nghịch với nhân tố này.
- Nhân tố 2 (ML2)
 - Biến có tải trọng cao:
 - * anxiety_level (0.867), depression (0.822), headache (0.835), mental_health_history (0.753), noise_level (0.747), study_load (0.741), future_career_concerns (0.840), peer_pressure (0.753), extracurricular_activities (0.763), bullying (0.858) có tải trọng rất cao, liên quan đến căng thẳng.
 - * breathing_problem (0.755) cũng góp phần, cho thấy yếu tố sinh lý khác.
 - Biến có tải trọng âm cao:
 - * self_esteem (-0.787), sleep_quality (-0.912), living_conditions (-0.790), safety (-0.907), basic_needs (-0.906), academic_performance (-0.909), teacher_student_relationship (-0.902), social_support (-0.719) có tải trọng âm mạnh, đại diện cho yếu tố tích cực.
 - Diễn giải:
 - * ML2 đại diện cho sự tương phản giữa các yếu tố gây căng thẳng (tích cực) và yếu tố bảo vệ chống căng thẳng (âm).
 - * Các biến như anxiety_level, depression, headache, bullying (tải trọng > 0.8) cho thấy nhân tố này phản ánh căng thẳng tâm lý, sinh lý, và môi trường.
 - * Các biến âm như self_esteem, sleep_quality, safety, academic_performance (tải trọng < -0.7) đại diện cho sức khỏe tinh thần, môi trường sống tốt, và hỗ trợ xã hội, là các yếu tố giảm căng thẳng

Giá trị Communalities:

```
print(fa.result$communalities)
```

```
##          anxiety_level          self_esteem
##          0.8775883          0.8525404
##    mental_health_history          depression
##          0.6670941          0.8646746
##          headache          blood_pressure
##          0.8288334          0.9950000
##          sleep_quality          breathing_problem
##          0.9217644          0.6144716
##          noise_level          living_conditions
##          0.7370965          0.7282124
##          safety          basic_needs
##          0.8599558          0.8611061
##    academic_performance          study_load
##          0.8658432          0.7328664
## teacher_student_relationship          future_career_concerns
##          0.8563694          0.9272373
##          social_support          peer_pressure
##          0.8972945          0.8603958
```

```
##      extracurricular_activities      bullying
##              0.8602667              0.9247457
```

Tổng quan

- Phạm vi giá trị: Từ 0.6144716 (breathing_problem) đến 0.9950000 (blood_pressure).
- Trung bình: Khoảng 0.84 (tính trung bình các giá trị), cho thấy tổng thể các nhân tố giải thích tốt dữ liệu.
- Nhận xét ban đầu: Hầu hết các biến có communalities cao (> 0.7), cho thấy mô hình 2 nhân tố phù hợp với phần lớn dữ liệu. Tuy nhiên, breathing_problem (0.614) là ngoại lệ đáng chú ý.
- Biến có communalities rất cao (> 0.9):
 - blood_pressure (0.995): Gần như toàn bộ phương sai được giải thích, phù hợp với tải trọng 0.991 trên ML1.
 - sleep_quality (0.922), future_career_concerns (0.927), social_support (0.897), bullying (0.925): Các biến này được giải thích rất tốt, phản ánh sự phân bố rõ ràng trong mô hình.
 - Nhận xét: Những biến này có cấu trúc nhân tố mạnh, phù hợp với phân tích trước (ML1 cho blood_pressure, ML2 cho các yếu tố căng thẳng/tích cực).
- Biến có communalities cao (0.8 - 0.9):
 - anxiety_level (0.878), self_esteem (0.853), depression (0.865), headache (0.829), safety (0.860), basic_needs (0.861), academic_performance (0.866), study_load (0.733), teacher_student_relationship (0.856), peer_pressure (0.860), extracurricular_activities (0.860).
 - Nhận xét: Các biến này cũng được giải thích tốt, phù hợp với diễn giải ML2 (căng thẳng vs. bảo vệ) và ML1 (huyết áp + căng thẳng nhẹ).
- Biến có communalities trung bình (0.6 - 0.7):
 - mental_health_history (0.667), breathing_problem (0.614), noise_level (0.737), living_conditions (0.728).
 - Nhận xét:
 - * breathing_problem (0.614) là biến có communality thấp nhất, cho thấy nó không được giải thích tốt bởi 2 nhân tố hiện tại. Điều này phù hợp với tải trọng 0.755 trên ML2, nhưng vẫn còn dư phương sai.
 - * mental_health_history (0.667) và noise_level (0.737) cũng có communality vừa phải, có thể cần thêm nhân tố để giải thích tốt hơn.

Tỷ lệ phương sai

```
print(fa.result$Vaccounted)
```

```
##              ML2      ML1
## SS loadings    12.8477920 3.8855594
## Proportion Var    0.6423896 0.1942780
## Cumulative Var    0.6423896 0.8366676
## Proportion Explained 0.7677955 0.2322045
## Cumulative Proportion 0.7677955 1.0000000
```

Nhận xét:

- SS loadings (Sum of Squared Loadings): ML2 đóng góp nhiều hơn gấp 3 lần so với ML1
- Proportion Var (Proportion of Total Variance Explained):
 - ML2 giải thích 64.24% tổng phương sai của dữ liệu.
 - ML1 bổ sung thêm 19.43%, đưa tổng phương sai được giải thích lên 83.67%.
- Proportion Explained (Proportion of Explained Variance):

- Đây là tỷ lệ phương sai được giải thích bởi mỗi nhân tố so với tổng phương sai đã được mô hình FA giải thích (83.67%).
- ML2 chiếm 76.78% trong số phương sai đã giải thích, trong khi ML1 chiếm 23.22%. Điều này cho thấy ML2 là nhân tố chính, trong khi ML1 đóng vai trò bổ sung.

```
cat("Thống kê chi-bình phương:", fa.result$STATISTIC, "\n")
```

```
## Thống kê chi-bình phương: 171.9397
```

```
cat("p-value:", fa.result$PVAL, "\n")
```

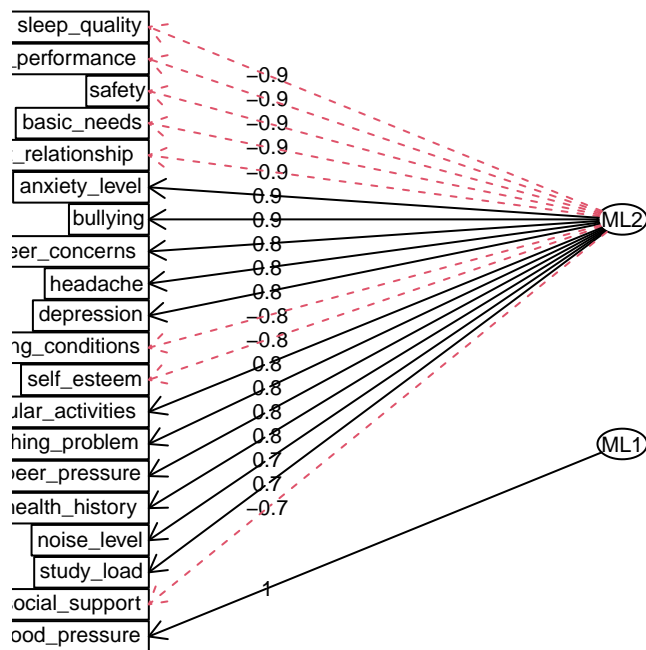
```
## p-value: 0.1167664
```

Nhận xét: Với p-value = 0.1167664 (> 0.05), mô hình FA hiện tại với 2 nhân tố (ML1 và ML2) được coi là phù hợp thống kê.

```
par(cex = 0.7)
```

```
fa.diagram(fa.result, main = "Biểu đồ nhân tố (Varimax Rotation)")
```

Biểu đồ nhân tố (Varimax Rotation)



Nhận

xét:

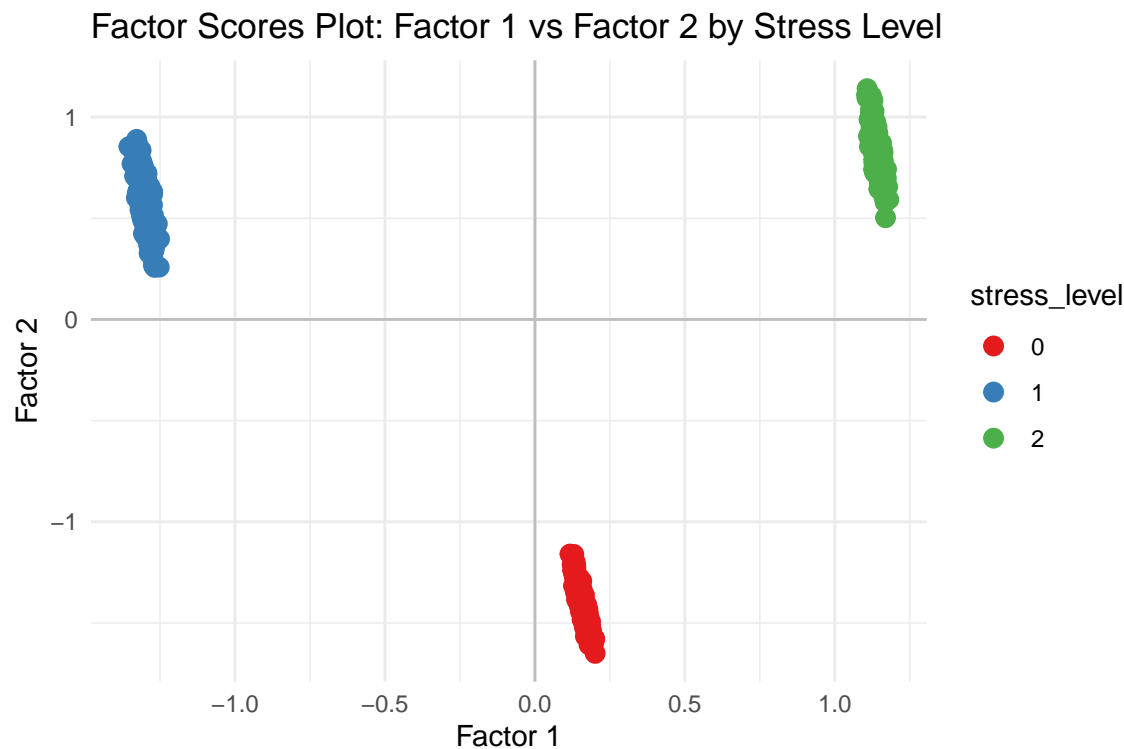
- Nhân tố ML1:
 - Có các biến mạnh mẽ liên quan như social_support (1.0), study_load (1.0), và peer_pressure (1.0), cho thấy nhân tố này có thể đại diện cho các yếu tố liên quan đến áp lực học tập và hỗ trợ xã hội.
 - Các biến khác như health_history (-0.7) và noise_level (-0.7) cũng có liên quan nhưng với cường độ thấp hơn và chiều ngược lại.
- Nhân tố ML2:
 - Liên quan mạnh mẽ đến các biến như sleep_quality (-0.9), academi_performance (-0.9), safety (-0.9), basic_needs (-0.9), teacher_student_relationship (-0.9), anxiety_level (0.9), bullying (0.9), future_career_concerns (0.8), headache (0.8), depression (-0.8), living_conditions (-0.8), self_esteem (0.8), extracurricular_activities (0.8), và solving_problem (0.7).

```

factor.scores <- fa.result$scores
factor.scores <- as.data.frame(factor.scores)
factor.scores$stress_level <- data.set.1$stress_level

# Vẽ biểu đồ phân tán (giả sử 3 nhân tố: MR1, MR2, MR3)
library(ggplot2)
ggplot(factor.scores, aes(x = ML1, y = ML2, color = stress_level)) +
  geom_point(size = 3) +
  labs(title = "Factor Scores Plot: Factor 1 vs Factor 2 by Stress Level",
       x = "Factor 1",
       y = "Factor 2") +
  geom_vline(xintercept = 0, color = "gray", linetype = "solid") +
  geom_hline(yintercept = 0, color = "gray", linetype = "solid") +
  theme_minimal() +
  scale_color_brewer(palette = "Set1")

```



Nhận xét: - Stress level 0 (đỏ): Tập trung ở vùng âm của Factor 1 (khoảng -1.0 đến -0.5) và gần 0 trên Factor 2. Điều này cho thấy nhóm này có xu hướng thấp trên Factor 1, có thể liên quan đến các yếu tố như áp lực học tập hoặc hỗ trợ xã hội (dựa trên biểu đồ trước).

- Stress level 1 (xanh dương): Tập trung ở vùng dương của Factor 2 (khoảng 0.5 đến 1.0) và gần 0 trên Factor 1. Nhóm này có xu hướng cao trên Factor 2, có thể liên quan đến các vấn đề sức khỏe tinh thần/thể chất như lo âu, trầm cảm.
- Stress level 2 (xanh lá): Tập trung ở vùng dương của cả Factor 1 (khoảng 0.5 đến 1.0) và Factor 2 (khoảng 0.5 đến 1.0). Nhóm này có xu hướng cao trên cả hai nhân tố, có thể chỉ ra sự kết hợp của áp lực học tập/hỗ trợ xã hội và các vấn đề sức khỏe.

KẾT LUẬN

Tổng kết chung:

- Cả PCA và EFA đều chỉ ra rằng bộ dữ liệu về mức độ căng thẳng của học sinh có thể được rút gọn hiệu quả thành hai yếu tố tiềm ẩn chính. Các yếu tố này không chỉ giúp tóm tắt thông tin phức tạp mà còn cho phép phân loại rõ ràng các mức độ căng thẳng khác nhau.
- Một nhân tố (chủ yếu là ML2) liên quan đến sự cân bằng giữa các vấn đề căng thẳng tâm lý, sinh lý, môi trường tiêu cực và các yếu tố bảo vệ như sức khỏe tinh thần tốt, môi trường sống an toàn và hỗ trợ xã hội.
- Nhân tố còn lại (chủ yếu là ML1) chủ yếu liên quan đến huyết áp và các khía cạnh của áp lực học tập và xã hội.
- Những phát hiện này cung cấp cái nhìn sâu sắc về các khía cạnh đa diện ảnh hưởng đến căng thẳng học đường, từ đó có thể định hướng các biện pháp can thiệp hiệu quả nhằm cải thiện sức khỏe tinh thần cho học sinh.