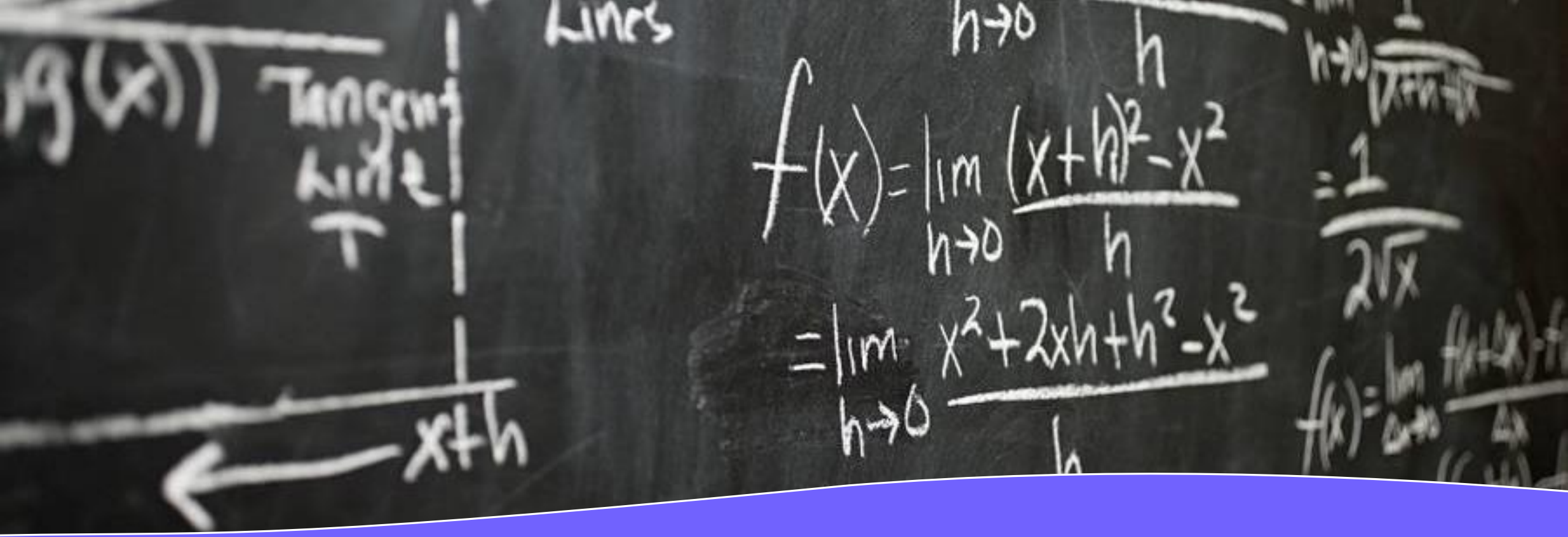


Faculty Courses Survey

Supervisor
Eng.Moataz

Team 18

Phase One

Proposal about Survey.

Proposal about Survey.

Introduction

Goals to be Accomplished

Survey Objectives

Participants & Survey design

Data analysis of the survey results

Timeline and important dates

ethics and confidentiality

Conclusion

Proposal

1) Introduction :

We are conducting a survey to gather feedback on the courses offered by our university. Your input is valuable to us as we strive to continuously improve our academic programs and ensure that they meet the needs of our students.

The purpose of this survey is to assess your satisfaction with the courses you have taken, as well as to identify areas where improvements can be made.

Proposal:

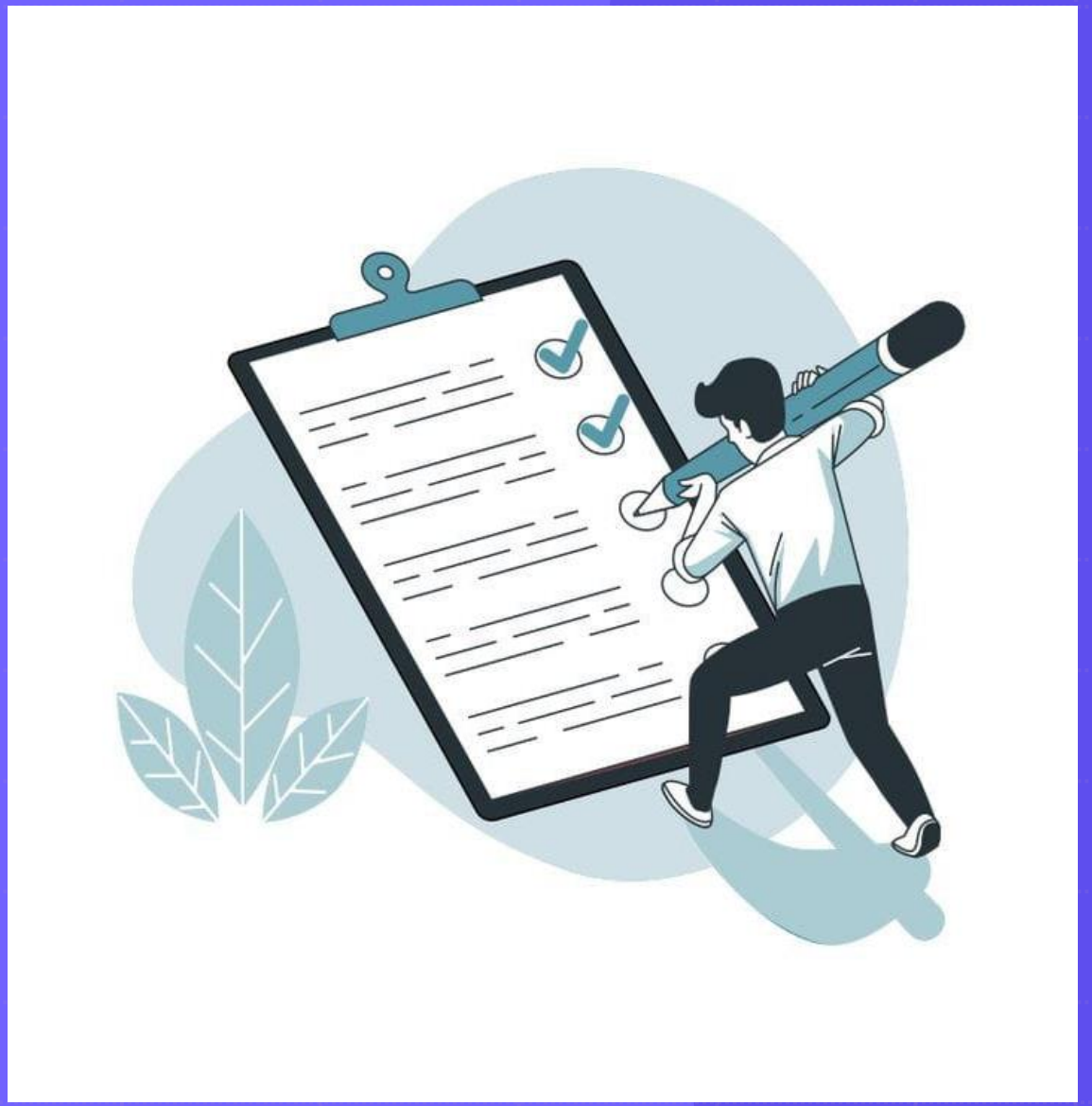
Goals to be Accomplished:

1. Identify areas of improvement: Surveys can help universities identify areas where they need to improve their courses. For example, if a large number of students report that they are struggling with a particular subject or topic, the university can take steps to provide additional resources or support
2. Evaluate course effectiveness: Surveys can help universities evaluate the effectiveness of their courses by asking students about their learning experiences and outcomes. This information can be used to make changes to the course content or delivery methods.
3. Gather feedback from students: Surveys provide an opportunity for students to provide feedback on their experiences with university courses. This feedback can be used to improve future courses and ensure that students have a positive learning experience.
4. Measure student satisfaction: Surveys can measure student satisfaction with university courses and overall academic experience. This information is important for universities to understand how they are meeting the needs and expectations of their students.
5. Inform decision-making: Survey results can inform decision-making at the university level, such as determining which courses should be offered in future semesters or which faculty members should be retained or promoted based on student feedback

Proposal

2) Survey Objectives:

- Assessing the faculty performance
- Gathering feedback on specific aspects of the courses
- Evaluating the quality and the effectiveness of the courses
- Measuring student’s satisfaction and study their interests
- Search for potential improvements
- Make the learning process a better experience.



Proposal

3) Participants

The participants are college students. The number of participants will be 200 college students.

They’re going to be convinced by telling them that it’s all for their own good and for the good of the college, and that it’s going to be good for them.

4) Survey design

Type of questions:

- 80 % close & 20 % open

-The survey will be online and will be given as a Google form through email or telegram

-Who have an apologize will take a survey written in a hard copy

Proposal

5) Data analysis of the survey results

-Once we have chosen our survey design, we will need to collect data. This might involve creating a survey questionnaire, recruiting participants, and administering the survey

- Clean and prepare data: Once we have collected data, we will need to clean and prepare it for analysis. This might involve checking for missing data, removing outliers, and coding open-ended responses.

- Choose analysis methods: There are many different methods you can use to analyze survey data, depending on the type of data we have collected and research question.

- Present your findings: Finally, you will need to present the findings concisely. This might involve creating tables, charts, and graphs to illustrate results and understand the meaning of statistical tests and identify any patterns or trends in data.

Proposal

6) Timeline and important dates:

Phase 1: we will hand over the proposal on 18/4/2023

Phase 2 & 3 : we will share our questionnaire between students and do sampling within 2 weeks (18/4/2023 : 2/5/2023)

Phase 4&5 : we will collect data & check expectations vs real within one week (2/5/2023 : 9/5/2023)

Phase 6 : we will apply data analysis within one week (9/5/2023 : 16/5/2023)

7) Budget of the project

The budget will be the costs of the software tools needed (computers, internet network, applications involved in the data collection, analysis and presentation).

Proposal

8) ethics and confidentiality

- For making sure the surveys are compliant with privacy laws. When collecting personal data, let survey takers know how you plan to use their data. Also, seeking the legal advice to comply with privacy laws and regulations.

- make responses anonymous and don't ask for sensitive information, whenever possible.

- The easiest way to protect confidentiality is to collect anonymous data.

- Anonymous data are data that are not connected to information that can identify the individual participant. And If there is no connection between the participant and their data, they are minimally risky.

- Not all studies can be anonymous. In the protocol, we need to justify why it is necessary to collect identifying information about an individual, include a list of identifiers that we will collect with the understanding that we will not collect more identifiers than needed, describe how this information will be used and how it will be collected, and describe what will be done to destroy this information once it is no longer needed.

Proposal

- If we collect personal information, we should consider how long we need it and why it needs to be stored. Don't store personal information for longer than necessary.

- We must always consider not only what will be best for our survey, but what is best for the participants who make the survey possible.

- All information gathered from the participants , they have the right to expect that it will not be divulged without their permission.

- In most cases, confidentiality can be ensured by using good data collection and storage practices. Make sure that all members of the study team understand not to discuss participants outside of the research context.

Proposal

9) Conclusion

- the proposed survey will provide valuable feedback about the quality of education provided to students.

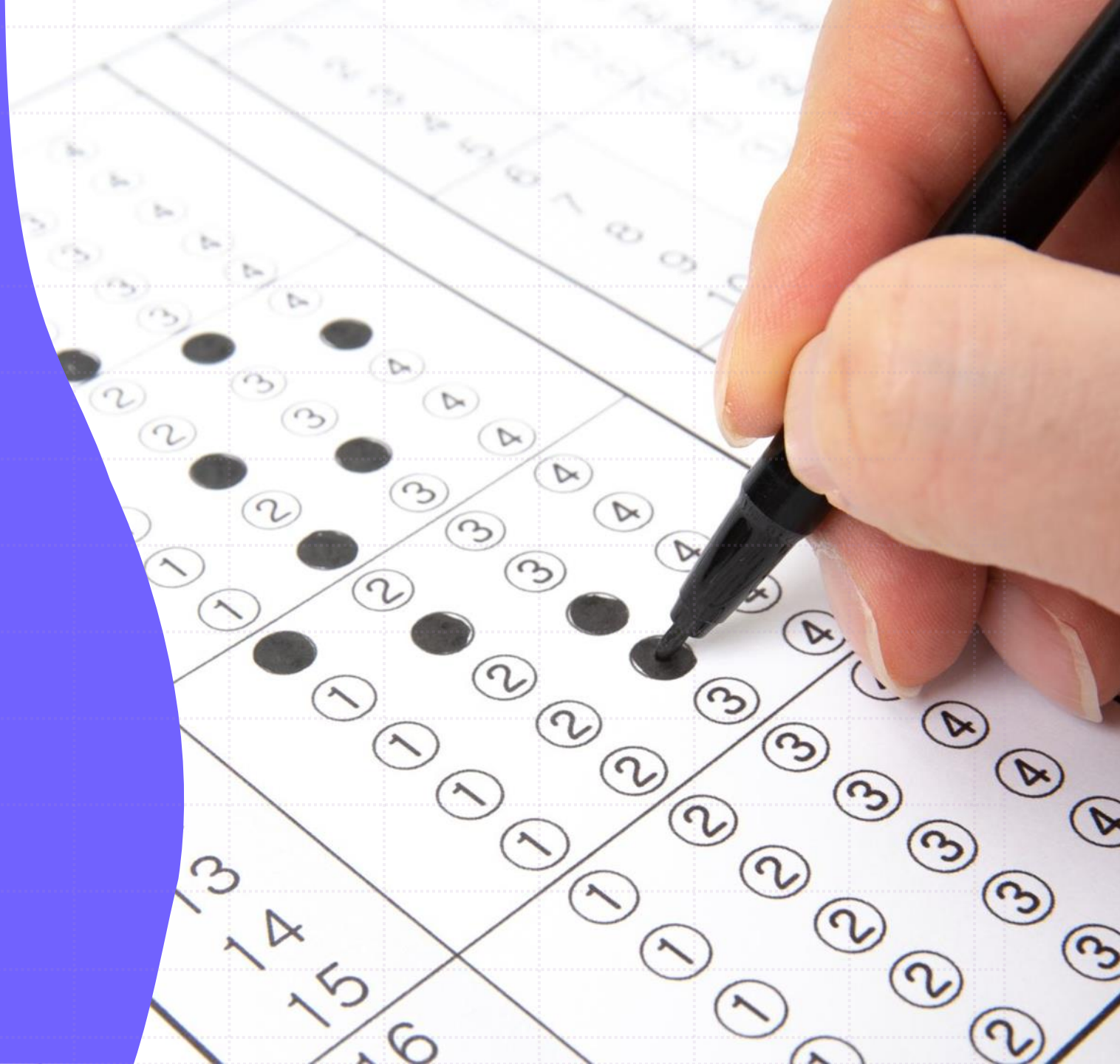
- The survey will cover a range of topics, including course content, teaching methods, organization and structure of the course, availability of course materials, effectiveness of assignments and exams.

-The survey results will be analyzed to identify areas where improvements can be made to the courses or teaching methods.

-The survey will be administered online and it will be voluntary, and all responses will be kept confidential

Phase Two

Questionnaire



Questionnaire

Introduction

Welcome to our Faculty Courses survey. We are interested in assessing the faculty performance and gathering feedback on the courses offered by the faculty.

Our goal in this survey is to get good attention on students' opinions on these courses and how to improve the benefits that students will get from them.

The survey will take **10 min** approximately.

Note: This data will not be published or shared outside the scope of survey

Demographic questions

1) What is your level?

   a) 1
   b) 2
   c) 3
   d) 4

2) Gender

   a) Male
   b) Female

3) What is your CGPA?

   a) 3.3: 4
   b) 3: 3.2
   c) 2: 2.9
   d) 1: 1.9
   e) Less than 1

Main survey questions

5) The provided material (Only by doctors or teaching assistants) is straight to the point and clear.
   a) Agree
   b) Maybe
   c) Disagree

6) The provided material (Only by doctors or teaching assistants) Include enough examples and practice materials.
   a) Agree
   b) Maybe
   c) Disagree

7) The provided material (Only by doctors or teaching assistants) is overall sufficient and reliable for studying.
   a) Agree
   b) Maybe
   c) Disagree

Response options

22) If you disagree with the previous question, How much do you think this set of courses qualifies you for the role.
   a) 0-30%
   b) 30-60%
   c) 60-90%
   d) 90-100% (agree with the previous question).

23) List the courses you think are crucial for your major and not provided by the faculty, if there's any

.....
.....
.....
.....
.....

24) Faculty courses provide you with soft skills like teamwork, communication, etc.
   a) Yes
   b) No

Optional questions

24) Faculty courses provide you with soft skills like teamwork, communication, etc.
   a) Yes
   b) No

25) If you agree to the previous question, list some of these skills.

.....
.....
.....
.....
.....

Closing

Thank you all for your valuable time!

Questionnaire

Introduction

Welcome to our Faculty Courses survey. We are interested in assessing the faculty performance and gathering feedback on the courses offered by the faculty.

Our goal in this survey is to get good attention on students' opinions on these courses and how to improve the benefits that students will get from them.

The survey will take **10 min** approximately.

Note: This data will not be published or shared outside the scope of survey

Questionnaire

Introduction

Note: This data will not be published or shared outside the scope of survey

We have ended our introduction with this important sentence which give respondents something of Reassurance and not be afraid to answer the survey.

Questionnaire

Introduction

Welcome to our Faculty Courses survey. We are interested in assessing the faculty performance and gathering feedback on the courses offered by the faculty.

Our goal in this survey is to get good attention on students' opinions on these courses and how to improve the benefits that students will get from them.

The survey will take **10 min** approximately.

Note: This data will not be published or shared outside the scope of survey

Demographic questions

1) What is your level?

   a) 1
   b) 2
   c) 3
   d) 4

2) Gender

   a) Male
   b) Female

3) What is your CGPA?

   a) 3.3: 4
   b) 3: 3.2
   c) 2: 2.9
   d) 1: 1.9
   e) Less than 1

Main survey questions

5) The provided material (Only by doctors or teaching assistants) is straight to the point and clear.
   a) Agree
   b) Maybe
   c) Disagree

6) The provided material (Only by doctors or teaching assistants) Include enough examples and practice materials.
   a) Agree
   b) Maybe
   c) Disagree

7) The provided material (Only by doctors or teaching assistants) is overall sufficient and reliable for studying.
   a) Agree
   b) Maybe
   c) Disagree

Response options

22) If you disagree with the previous question, How much do you think this set of courses qualifies you for the role.
   a) 0-30%
   b) 30-60%
   c) 60-90%
   d) 90-100% (agree with the previous question).

23) List the courses you think are crucial for your major and not provided by the faculty, if there's any

.....
.....
.....
.....
.....

24) Faculty courses provide you with soft skills like teamwork, communication, etc.
   a) Yes
   b) No

Optional questions

24) Faculty courses provide you with soft skills like teamwork, communication, etc.
   a) Yes
   b) No

25) If you agree to the previous question, list some of these skills.

.....
.....
.....
.....
.....

Closing

Thank you all for your valuable time!

Questionnaire

Demographic questions

1) What is your level?

   a) 1
   b) 2
   c) 3
   d) 4

2) Gender

   a) Male
   b) Female

3) What is your CGPA?

   a) 3.3: 4
   b) 3: 3.2
   c) 2: 2.9
   d) 1: 1.9
   e) Less than 1

Questionnaire

Introduction

Welcome to our Faculty Courses survey. We are interested in assessing the faculty performance and gathering feedback on the courses offered by the faculty.

Our goal in this survey is to get good attention on students' opinions on these courses and how to improve the benefits that students will get from them.

The survey will take **10 min** approximately.

Note: This data will not be published or shared outside the scope of survey

Demographic questions

1) What is your level?

   a) 1
   b) 2
   c) 3
   d) 4

2) Gender

   a) Male
   b) Female

3) What is your CGPA?

   a) 3.3: 4
   b) 3: 3.2
   c) 2: 2.9
   d) 1: 1.9
   e) Less than 1

Main survey questions

5) The provided material (Only by doctors or teaching assistants) is straight to the point and clear.
   a) Agree
   b) Maybe
   c) Disagree

6) The provided material (Only by doctors or teaching assistants) Include enough examples and practice materials.
   a) Agree
   b) Maybe
   c) Disagree

7) The provided material (Only by doctors or teaching assistants) is overall sufficient and reliable for studying.
   a) Agree
   b) Maybe
   c) Disagree

Response options

22) If you disagree with the previous question, How much do you think this set of courses qualifies you for the role.
   a) 0-30%
   b) 30-60%
   c) 60-90%
   d) 90-100% (agree with the previous question).

23) List the courses you think are crucial for your major and not provided by the faculty, if there's any

.....
.....
.....
.....
.....

24) Faculty courses provide you with soft skills like teamwork, communication, etc.
   a) Yes
   b) No

Optional questions

24) Faculty courses provide you with soft skills like teamwork, communication, etc.
   a) Yes
   b) No

25) If you agree to the previous question, list some of these skills.

.....
.....
.....
.....
.....

Closing

Thank you all for your valuable time!

Questionnaire

Main survey questions

5) The provided material (Only by doctors or teaching assistants) is straight to the point and clear.
    a) Agree
    b) Maybe
    c) Disagree

6) The provided material (Only by doctors or teaching assistants) Include enough examples and practice materials.
    a) Agree
    b) Maybe
    c) Disagree

7) The provided material (Only by doctors or teaching assistants) is overall sufficient and reliable for studying.
    a) Agree
    b) Maybe
    c) Disagree

Questionnaire

- Questions are divided as following :

80 % closed ended questions.

20% open ended questions.

- Questions are divided into 5 sections as follows:

1. Questions related to course content.
2. Questions related to course relevancy to the future professional role.
3. Questions related to the teaching staff.
4. Questions related to campus environment.
5. Questions related to exams & assignments.

Main survey questions

5) The provided material (Only by doctors or teaching assistants) is straight to the point and clear.
   a) Agree
   b) Maybe
   c) Disagree

6) The provided material (Only by doctors or teaching assistants) Include enough examples and practice materials.
   a) Agree
   b) Maybe
   c) Disagree

7) The provided material (Only by doctors or teaching assistants) is overall sufficient and reliable for studying.
   a) Agree
   b) Maybe
   c) Disagree

Questionnaire

Introduction

Welcome to our Faculty Courses survey. We are interested in assessing the faculty performance and gathering feedback on the courses offered by the faculty.

Our goal in this survey is to get good attention on students' opinions on these courses and how to improve the benefits that students will get from them.

The survey will take **10 min** approximately.

Note: This data will not be published or shared outside the scope of survey

Demographic questions

1) What is your level?

   a) 1
   b) 2
   c) 3
   d) 4

2) Gender

   a) Male
   b) Female

3) What is your CGPA?

   a) 3.3: 4
   b) 3: 3.2
   c) 2: 2.9
   d) 1: 1.9
   e) Less than 1

Main survey questions

5) The provided material (Only by doctors or teaching assistants) is straight to the point and clear.
   a) Agree
   b) Maybe
   c) Disagree

6) The provided material (Only by doctors or teaching assistants) Include enough examples and practice materials.
   a) Agree
   b) Maybe
   c) Disagree

7) The provided material (Only by doctors or teaching assistants) is overall sufficient and reliable for studying.
   a) Agree
   b) Maybe
   c) Disagree

Response options

22) If you disagree with the previous question, How much do you think this set of courses qualifies you for the role.
   a) 0-30%
   b) 30-60%
   c) 60-90%
   d) 90-100% (agree with the previous question).

23) List the courses you think are crucial for your major and not provided by the faculty, if there's any

.....
.....
.....
.....
.....

24) Faculty courses provide you with soft skills like teamwork, communication, etc.
   a) Yes
   b) No

Optional questions

24) Faculty courses provide you with soft skills like teamwork, communication, etc.
   a) Yes
   b) No

25) If you agree to the previous question, list some of these skills.

.....
.....
.....
.....
.....

Closing

Thank you all for your valuable time!

Questionnaire

Response options

22) If you disagree with the previous question, How much do you think this set of courses qualifies you for the role.

a) 0-30%
b) 30-60%
c) 60-90%
d) 90-100% (agree with the previous question).

23) List the courses you think are crucial for your major and not provided by the faculty, if there's any

.....
.....
.....
.....
.....

24) Faculty courses provide you with soft skills like teamwork, communication, etc.

a) Yes
b) No

Questionnaire

We have provided options in the questions by providing :

- Rates
- MCQ's
- Open ended questions for the respondents to express their thoughts

Response options

22) If you disagree with the previous question, How much do you think this set of courses qualifies you for the role.
   a) 0-30%
   b) 30-60%
   c) 60-90%
   d) 90-100% (agree with the previous question).

23) List the courses you think are crucial for your major and not provided by the faculty, if there's any
.....
.....
.....
.....
.....

24) Faculty courses provide you with soft skills like teamwork, communication, etc.
   a) Yes
   b) No

Questionnaire

Introduction

Welcome to our Faculty Courses survey. We are interested in assessing the faculty performance and gathering feedback on the courses offered by the faculty.

Our goal in this survey is to get good attention on students' opinions on these courses and how to improve the benefits that students will get from them.

The survey will take **10 min** approximately.

Note: This data will not be published or shared outside the scope of survey

Demographic questions

1) What is your level?

   a) 1
   b) 2
   c) 3
   d) 4

2) Gender

   a) Male
   b) Female

3) What is your CGPA?

   a) 3.3: 4
   b) 3: 3.2
   c) 2: 2.9
   d) 1: 1.9
   e) Less than 1

Main survey questions

5) The provided material (Only by doctors or teaching assistants) is straight to the point and clear.
   a) Agree
   b) Maybe
   c) Disagree

6) The provided material (Only by doctors or teaching assistants) Include enough examples and practice materials.
   a) Agree
   b) Maybe
   c) Disagree

7) The provided material (Only by doctors or teaching assistants) is overall sufficient and reliable for studying.
   a) Agree
   b) Maybe
   c) Disagree

Response options

22) If you disagree with the previous question, How much do you think this set of courses qualifies you for the role.
   a) 0-30%
   b) 30-60%
   c) 60-90%
   d) 90-100% (agree with the previous question).

23) List the courses you think are crucial for your major and not provided by the faculty, if there's any

.....
.....
.....
.....
.....

24) Faculty courses provide you with soft skills like teamwork, communication, etc.
   a) Yes
   b) No

Optional questions

24) Faculty courses provide you with soft skills like teamwork, communication, etc.
   a) Yes
   b) No

25) If you agree to the previous question, list some of these skills.

.....
.....
.....
.....
.....

Closing

Thank you all for your valuable time!

Questionnaire

Optional questions

24) Faculty courses provide you with soft skills like teamwork, communication, etc.

a) Yes
b) No

25) If you agree to the previous question, list some of these skills.

.....

.....

.....

.....

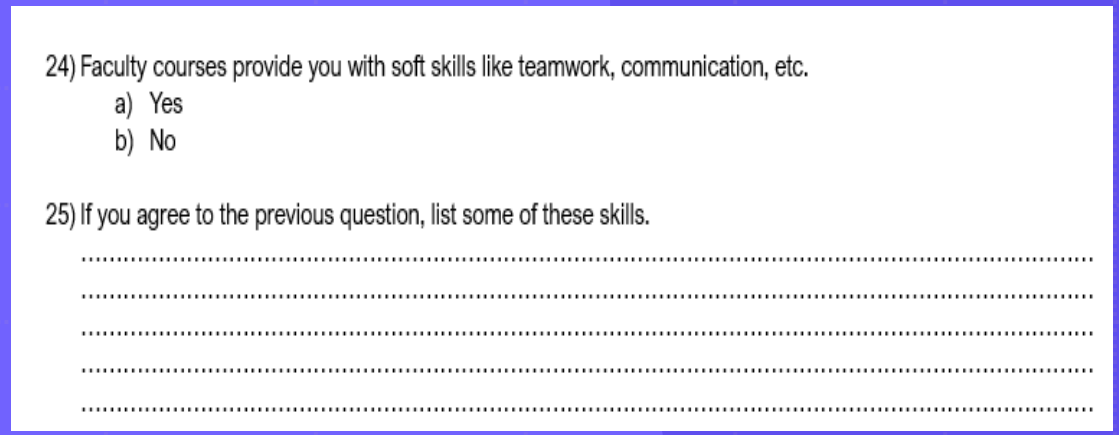
.....

Questionnaire

We have provided optional questions as the follows:

- If a respondent agrees or disagrees with a question, they have all the freedom to explain their point of view and express their thoughts

Optional questions



24) Faculty courses provide you with soft skills like teamwork, communication, etc.

a) Yes
b) No

25) If you agree to the previous question, list some of these skills.

.....
.....
.....
.....
.....

Questionnaire

Introduction

Welcome to our Faculty Courses survey. We are interested in assessing the faculty performance and gathering feedback on the courses offered by the faculty.

Our goal in this survey is to get good attention on students' opinions on these courses and how to improve the benefits that students will get from them.

The survey will take **10 min** approximately.

Note: This data will not be published or shared outside the scope of survey

Demographic questions

1) What is your level?

   a) 1
   b) 2
   c) 3
   d) 4

2) Gender

   a) Male
   b) Female

3) What is your CGPA?

   a) 3.3: 4
   b) 3: 3.2
   c) 2: 2.9
   d) 1: 1.9
   e) Less than 1

Main survey questions

5) The provided material (Only by doctors or teaching assistants) is straight to the point and clear.
   a) Agree
   b) Maybe
   c) Disagree

6) The provided material (Only by doctors or teaching assistants) Include enough examples and practice materials.
   a) Agree
   b) Maybe
   c) Disagree

7) The provided material (Only by doctors or teaching assistants) is overall sufficient and reliable for studying.
   a) Agree
   b) Maybe
   c) Disagree

Response options

22) If you disagree with the previous question, How much do you think this set of courses qualifies you for the role.
   a) 0-30%
   b) 30-60%
   c) 60-90%
   d) 90-100% (agree with the previous question).

23) List the courses you think are crucial for your major and not provided by the faculty, if there's any

.....
.....
.....
.....
.....

24) Faculty courses provide you with soft skills like teamwork, communication, etc.
   a) Yes
   b) No

Optional questions

24) Faculty courses provide you with soft skills like teamwork, communication, etc.
   a) Yes
   b) No

25) If you agree to the previous question, list some of these skills.

.....
.....
.....
.....
.....

Closing

Thank you all for your valuable time!

Questionnaire

Closing

Thank you all for your valuable time!

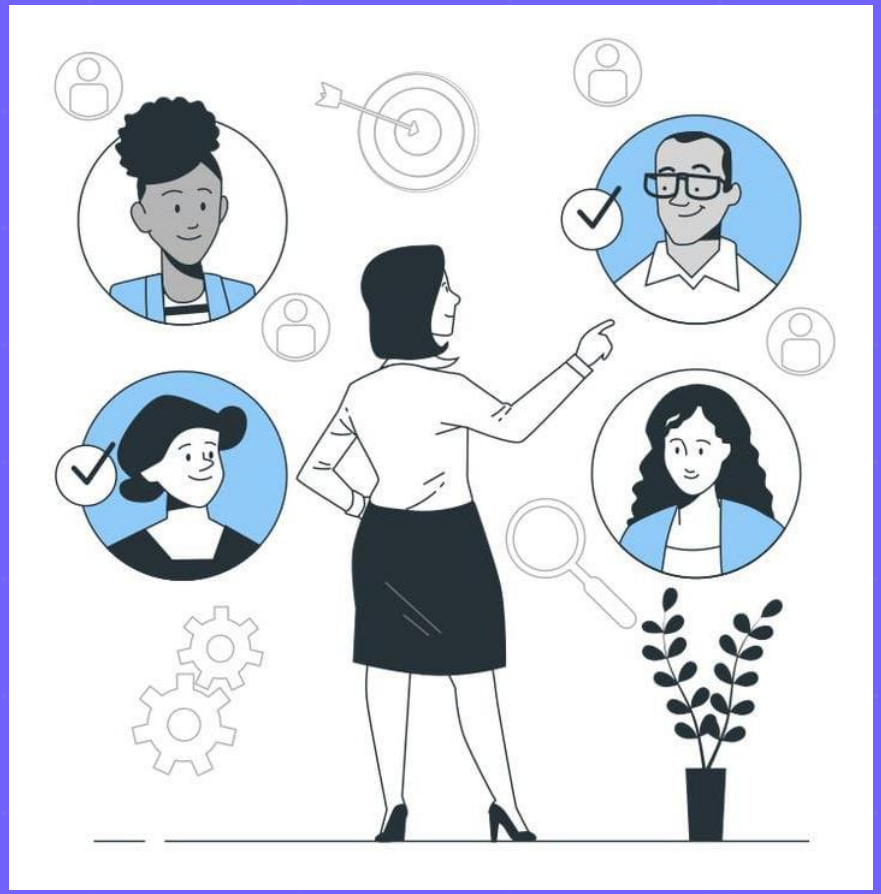
We have ended our questionnaire with this sentence to give the respondents appreciation for filling our questionnaire.

Phase Three

Sampling & Pretests.

Sampling & Pretests.

Sampling phase



Sampling & Pretests.

Sampling phase :

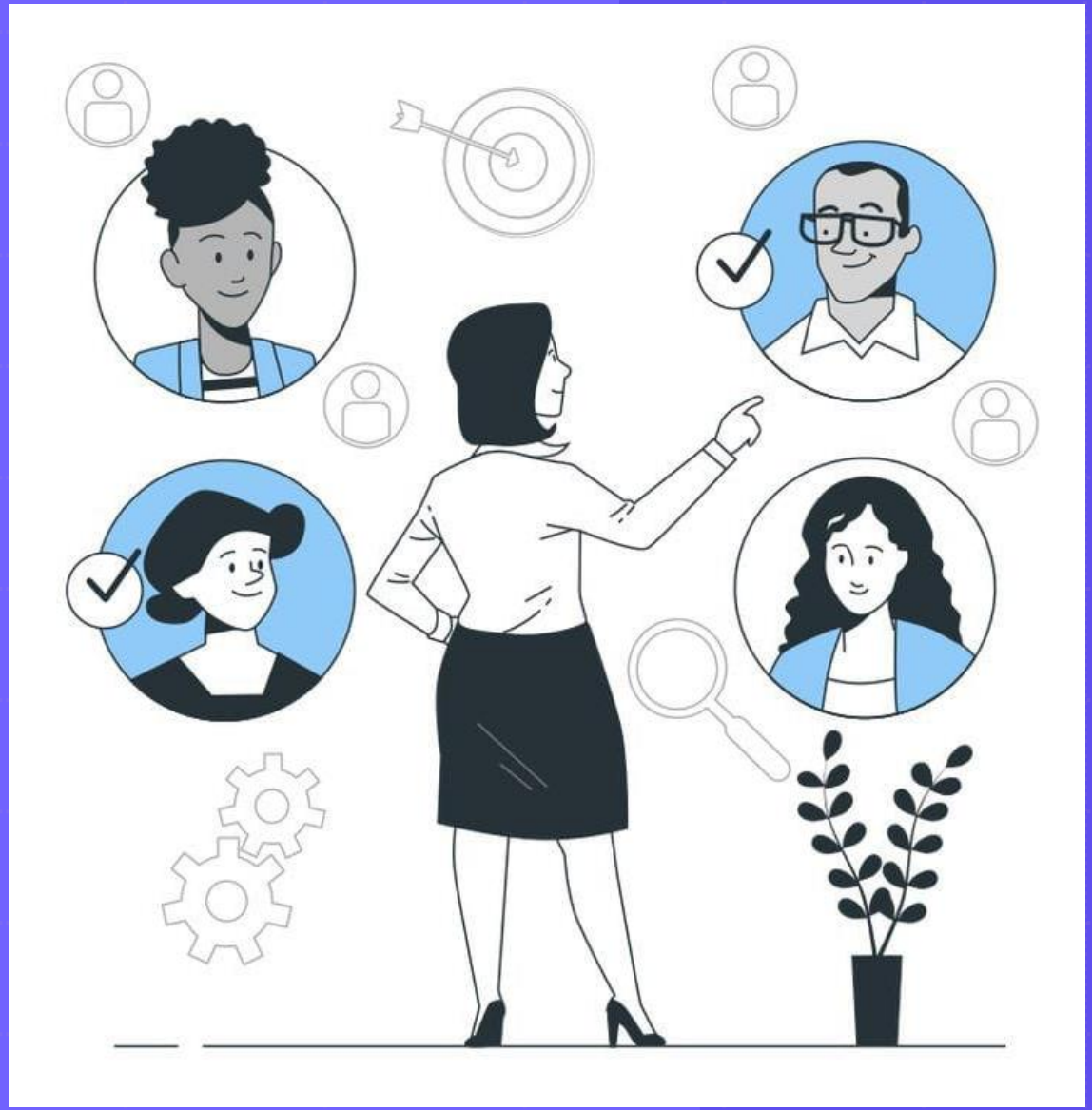
Our population are student of our college

We will take a sample of about dent to gather information about the population for pretest

The goal of sampling is to obtain a representative sample that accurately reflects the characteristics of the larger population being studied.

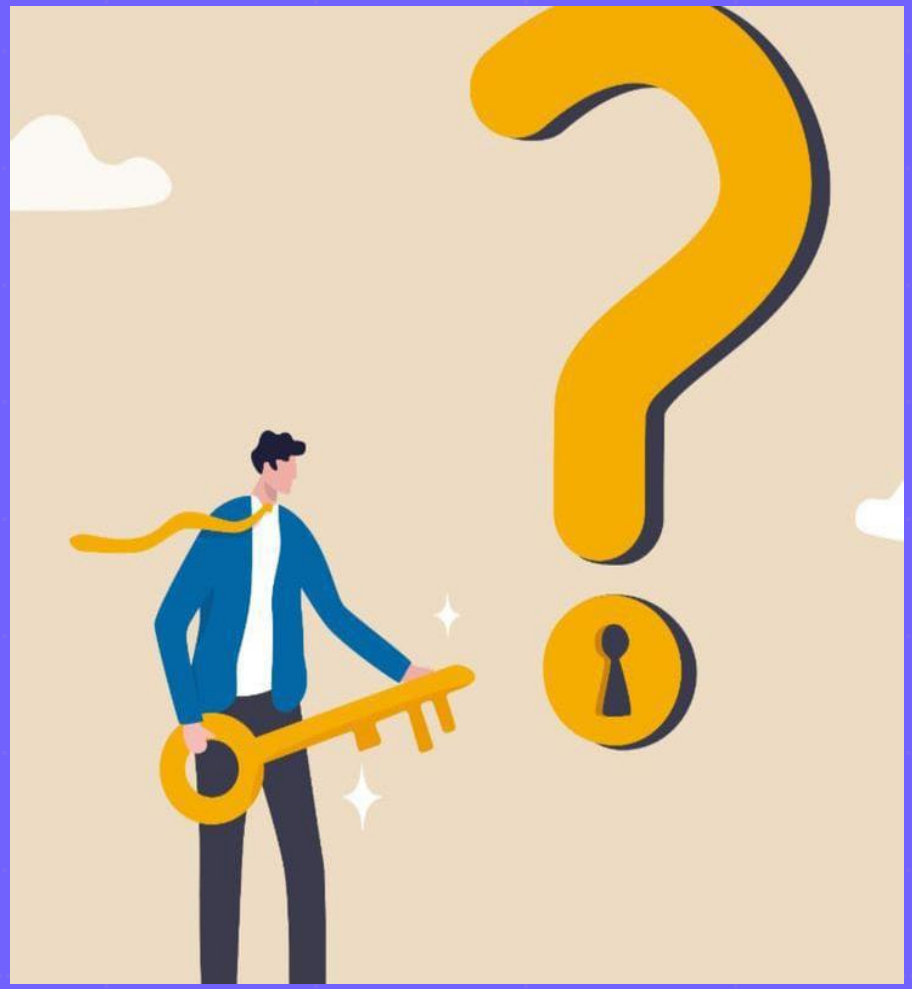
This could be any group of people or entities that share some common characteristic, such as gpa,level

Our sampling method is cluster sampling



Sampling & Pretests.

Pretest phase

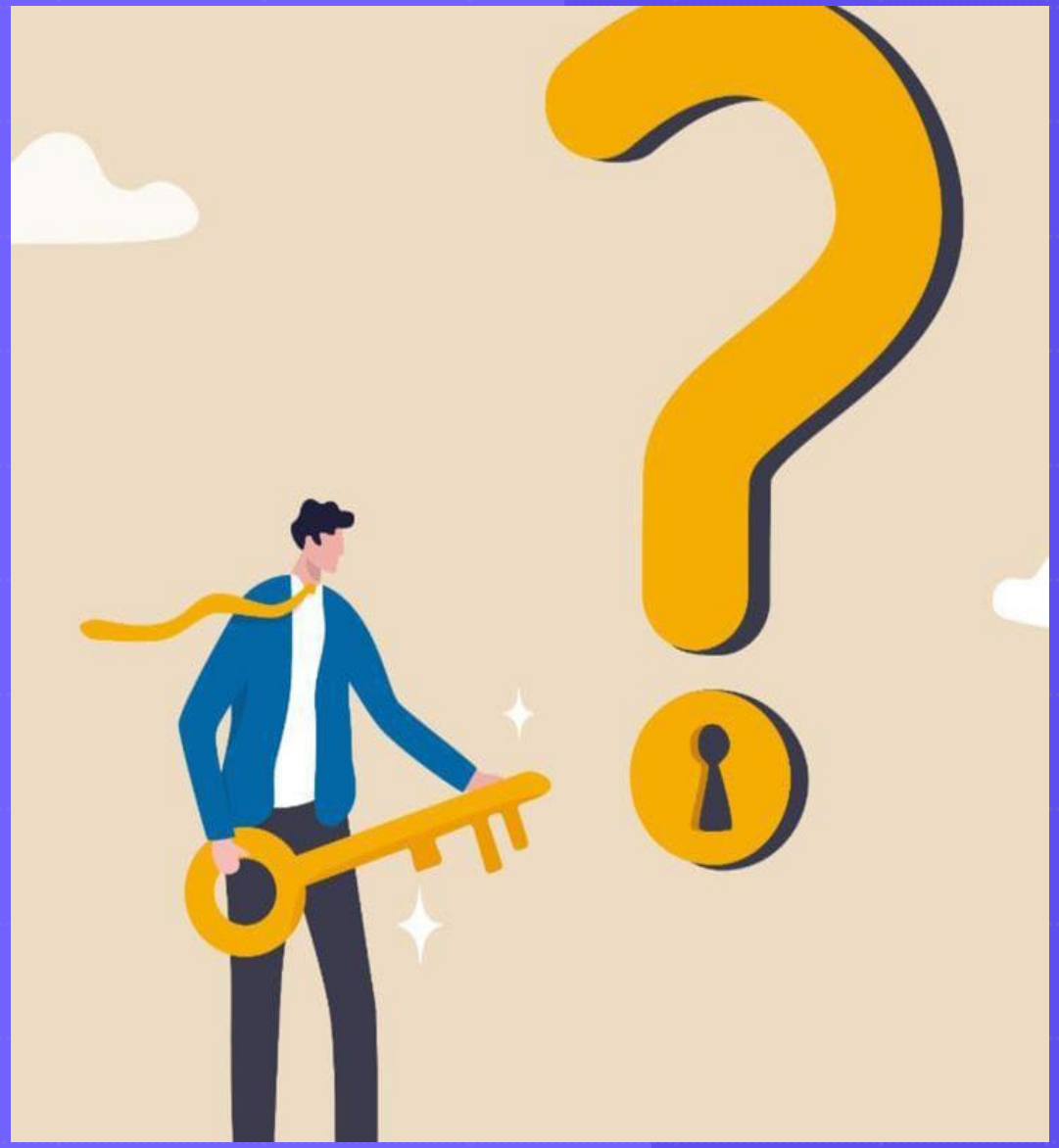


Sampling & Pretests.

Pretest phase:

The pretest is typically conducted on a small sample of individuals (about 40) who are similar to the target population in terms of demographic characteristics (gpa, level).

The pretest phase is important because it helps to ensure that the survey instrument is clear, relevant, and unbiased, which in turn improves the quality of the data collected and the accuracy of the survey results. By identifying and correcting problems with the questionnaire before it is administered to the target population, the pretest phase can save time and resources and improve the overall success of the survey project.



Sampling & Pretests

Comprehension:

Campus word used in questions may be not understandable with all respondent

We may explain this word : the building of the college

Logic and flow :

There are 5 questions about halls & labs and may be out of our scope “faculty courses”

Acceptability

We add id attribute to ensure that every respondent will send his answer once

But we found non acceptable answers

May be it touch on sensitive subject

Sampling & Pretests

Length and adherence:

The questionnaire experience may be effortful and most respondents find it's difficult to make it through to the end without losing interest and focus.

We can divide the questionnaire into 2 parts

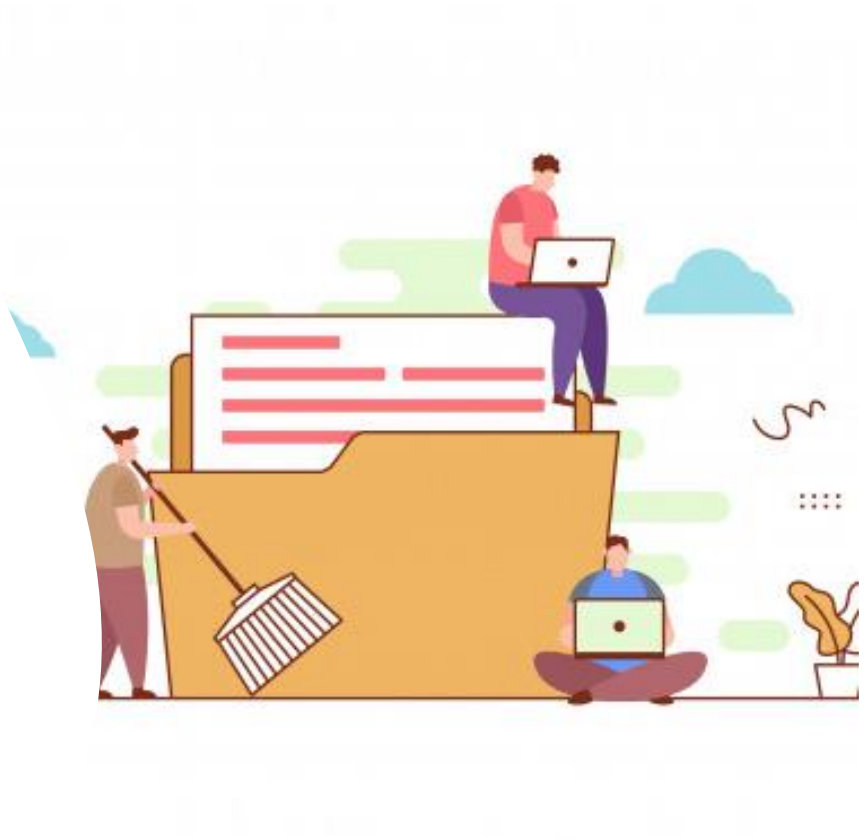
Part1: about courses

Part2: about staff

Introduction and gaining consent

Phase 4

Doing the pre-analysis on the data



Sample of the data before any cleaning

Out[39]:

	level	gender	CGPA	theoretical part	material clearness	enough practice examples	material sufficiency	ideas to improve materials	timeframe suitability	CC opnions	...	learning environment can affect performance	rooms/Labs are clean	rooms/Labs are big	rooms/ have t
0	2	Female	3.3 : 4	well enough to pass.	Maybe	Agree	Maybe	It should be more simple and comprehensive	Maybe	It is too long for the semester.	...	Yes	No	No	
1	3	Female	3.3 : 4	Very well.	Maybe	Maybe	Disagree	To have the material up to date where we can l...	Agree.	I agree with the previous question.	...	Yes	No	No	
2	2	Female	3.3 : 4	well enough to pass.	Agree	Maybe	Maybe	Need to cover more topics or clearer informati...	Agree.	I agree with the previous question.	...	Yes	No	No	
3	3	Male	3.3 : 4	Very well.	Agree	Agree	Disagree	No	Maybe	I agree with the previous question.	...	Yes	No	No	
4	3	Female	3.3 : 4	Very well.	Maybe	Maybe	Maybe	To be more detailed than it is to depend on it	Maybe	I agree with the previous question.	...	Yes	Yes	No	

Checking data type of each column , As we see all columns are from type object , except level column is int

```
# Check the data types of the columns
survey_data.dtypes
```

```
Timestamp                                                                 object
Faculty                                                                   object
 ID                                                                       object
department                                                                object
 level                                                                     int64
gender                                                                    object
CGPA                                                                      object
understanding the theoretical part of courses                             object
whether material are clear or not                                         object
whether material has enough  practice examples                            object
whether material  is overall sufficient                                   object
ideas to improve the quality of the provided material                     object
whether Course content is suitable for the semester timeframe.            object
If you disagree with the previous question, what do you think about the content?    object
 if assignments/projects are directly related to the course content       object
if assignments/projects successfully measure your acquired knowledge      object
if the project discussions provide helpful notes and evaluation.          object
if Attending lectures  improves level of understanding of the course      object
if Labs/Classes improves level of understanding of the course             object
suggestions for enhancing  learning experience inside campus             object
doctors cooporations                                                      object
 teaching assistants cooporations                                         object
if doctors and teaching assistants use interactive teaching techniques    object
if interactive techniques can improve performance                         object
if  the faculty provides the best set of courses to qualify you for your desired  role    object
How much do you think this set of courses qualifies you for the role.     object
List the courses you think are crucial for your major and not provided by the faculty    object
Faculty courses provide you with soft skills                              object
 list some of of soft skills provided by college                          object
whether learning environment can affect  performance                      object
whether lecture rooms/Labs are clean and organized.                       object
if lecture rooms/Labs are big enough for the number of students.          object
```

Checking and handling missing values

In [7]:
```
# Check for missing values
survey_data.isnull().sum()
```

Out[7]:

Timestamp	0
Faculty	0
ID	50
department	38
level	0
gender	0
CGPA	0
understanding the theoretical part of courses	0
whether material are clear or not	0
whether material has enough practice examples	0
whether material is overall sufficient	0
ideas to improve the quality of the provided material	0
whether Course content is suitable for the semester timeframe.	0
If you disagree with the previous question, what do you think about the content?	0
if assignments/projects are directly related to the course content	0
if assignments/projects successfully measure your acquired knowledge	0
if the project discussions provide helpful notes and evaluation.	0
if Attending lectures improves level of understanding of the course	0
if Labs/Classes improves level of understanding of the course	0
suggestions for enhancing learning experience inside campus	2
doctors cooporations	0
teaching assistants cooporations	0
if doctors and teaching assistants use interactive teaching techniques	0
if interactive techniques can improve performance	0
if the faculty provides the best set of courses to qualify you for your desired role	0
How much do you think this set of courses qualifies you for the role.	0
List the courses you think are crucial for your major and not provided by the faculty	2
Faculty courses provide you with soft skills	0
list some of of soft skills provided by college	45
whether learning environment can affect performance	0
whether lecture rooms/Labs are clean and organized.	0
if lecture rooms/Labs are big enough for the number of students.	0
whether lecture rooms/Labs have suitable and clear board/screens to present the materials	0
The halls are equipped for gaps and breaks.	0
suggestions to improve experience inside campus	2
Rating of overall satisfaction with faculty courses.	0

```
survey_data.fillna(value=0, inplace=True)
```

Handling duplicates by dropping them and Handling outliers using inter quartile range if there is any

```
In [9]: # Identify and remove duplicate responses
        survey_data.drop_duplicates(inplace=True)
```

```
# Identify outliers using IQR
Q1 = survey_data.quantile(0.25)
Q3 = survey_data.quantile(0.75)
IQR = Q3 - Q1

# Remove outliers
survey_data = survey_data[~((survey_data < (Q1 - 1.5 * IQR)) | (survey_data > (Q3 + 1.5 * IQR))).any(axis=1)]
```

rename columns using excel

Removing careless responses

```
survey_data['ideas to improve materials'] = survey_data['ideas to improve materials'].astype(str)
survey_data['suggestions for BL'] = survey_data['suggestions for BL'].astype(str)
survey_data['list of IC'] = survey_data['list of IC'].astype(str)
survey_data['soft skills suggistions'] = survey_data['soft skills suggistions'].astype(str)
survey_data['suggestions for better exp'] = survey_data['suggestions for better exp'].astype(str)
```

this lines used to convert the responses of questions to strings

```
#col -> ideas to improve materials

# define list of unimportant words
unimportant_words = ['other', 'no', 'nothing', 'i dont know', '...', '.', '-', '@','$','/']
```

This line defines a list of words and characters that are considered unimportant and should be filtered out from the text data.

```
# define function to check if sentence contains unimportant words or is one character long
def check_sentence(sentence):
    if any(word in sentence.lower() for word in unimportant_words) or len(sentence) == 1:
        return np.nan
    else:
        return sentence
```

The previous code defines a function called check_sentence() that takes a sentence as input and checks if it contains any of the unimportant words or characters, or if it is only one character long. If the sentence contains any of the unimportant words or characters or is only one character long, the function returns NaN. Otherwise, it returns the original sentence

```
# apply the check_sentence function to the "suggestions to improve experience inside campus" column
#to replace unimportant sentences with NaN
survey_data['ideas to improve materials'] = survey_data['ideas to improve materials'].apply(lambda x: check_sentence(x)
                                                                                            if isinstance(x, str) else x)
```

This line applies the check_sentence() function to the "ideas to improve materials" column of the survey_data DataFrame using the apply() method.
The lambda function is used to check if the input value is a string before applying the check_sentence() function. If the input value is not a string (e.g., it is NaN or a non-string value), the lambda function returns
the input value unchanged. If the input value is a string, the lambda function applies the check_sentence() function to the string and returns the result.

```
survey_data['ideas to improve materials']
```

```
0           It should be more simple and comprehensive
1                                                  NaN
2     Need to cover more topics or clearer informati...
3                                                  NaN
4        To be more detailed than it is to depend on it
                            ...
94                                                 NaN
95                                                 NaN
96                                                 NaN
97                                                 NaN
98                                  Books and work shop
Name: ideas to improve materials, Length: 99, dtype: object
```

removing very small and large response text:

```
# Specify the column name
column_name = 'suggestions for BL'
# Loop through each row in the column and calculate the length of the text
text_lengths_b = []
text_lengths_l = []
for index ,row in survey_data.iterrows():
    text = row[column_name]
    if(len(text)>200):
        text_length_b = len(text)
        text_lengths_b.append(text_length_b)
        row[column_name]=None
    elif(len(text)< 3):
        text_length_l = len(text)
        text_lengths_l.append(text_length_l)
        row[column_name]=None

# Print the list of text with high length
print(text_lengths_b)
# Print the list of text with low length
print(text_lengths_l)
survey_data['suggestions for BL']
```

1.Defines a column name as suggestions for BL(better learning)
2.Loops through each row in the suggestions for better exp column of the survey_data DataFrame and calculates the length of the text in that column for each row.
3.For each row, if the length of the text is greater than 200, it appends the length to a list called text_lengths_b and sets the value of the cell to None. If the length of the text is less than 3, it appends the length to a list called text_lengths_l and also sets the value of the cell to None.
4.Prints the lists of text lengths for high and low length texts.

A look on the data after cleaning it

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	level	gender	CGPA	theoretica	material cl	enough pr	material su	ideas to im	timeframe	CC opnion	assignmen	assignmen	project dis	Attending	Labs/Class	suggestion	doctors co	teaching a	interactive	interactive	courses re	courses qu	list of IC
2	2	Female	303:04:00	well enoug	Maybe	Agree	Maybe	It should b	Maybe	It is too lo	Disagree	Maybe	Maybe	Yes	Yes	By working	Always co	Very coop	Some of th	Yes	Yes	90-100% (a	Null
3	3	Female	303:04:00	Very well0	Maybe	Maybe	Disagree	0	Agree0	I agree wit	Agree	Maybe	Maybe	Yes	Yes	0	Very coop	Very coop	Some of th	Yes	No	30-60%	0
4	2	Female	303:04:00	well enoug	Agree	Maybe	Maybe	Need to co	Agree0	I agree wit	Agree	Disagree	Maybe	Yes	Yes	Classes ne	Very coop	Very coop	Agree	No	Yes	90-100% (a	0
5	3	Male	303:04:00	Very well0	Agree	Agree	Disagree	0	Maybe	I agree wit	Agree	Maybe	Maybe	No	Yes	The faculty	Very coop	Never coo	Some of th	Yes	Yes	90-100% (a	Visualizat
6	3	Female	303:04:00	Very well0	Maybe	Maybe	Maybe	To be mor	Maybe	I agree wit	Disagree	Maybe	Disagree	Yes	No	0	Always co	Always co	Some of th	Yes	No	30-60%	0
7	2	Female	303:04:00	well enoug	Agree	Maybe	Disagree	take stude	Disagree0	It is too lo	Agree	Maybe	Agree	Yes	Yes	0	Always co	Always co	Some of th	Yes	Yes	90-100% (a	0
8	3	Female	02:02.9	Very well0	Maybe	Maybe	Maybe	Provide m	Maybe	It is too lo	Maybe	Disagree	Disagree	Yes	Yes	More help	Never coo	Very coop	Some of th	Yes	No	0-30%	Design pa
9	3	Male	303:04:00	Very well0	Maybe	Maybe	Disagree	Provide m	Maybe	It is too lo	Maybe	Maybe	Disagree	No	Yes	Provide m	Very coop	Always co	Agree	Yes	No	30-60%	Neural ne
10	3	Male	303:04:00	Very well0	Maybe	Maybe	Disagree	We should	Agree0	I agree wit	Agree	Maybe	Disagree	Yes	No	0	Very coop	Always co	Agree	Yes	No	60-90%	0
11	3	Male	02:02.9	well enoug	Maybe	Disagree	Maybe	I think that	Maybe	It is too lo	Maybe	Maybe	Agree	Yes	Yes	0	Very coop	Always co	Some of th	Yes	No	0-30%	Data
12	2	Female	02:02.9	well enoug	Maybe	Maybe	Maybe	attach it w	Disagree0	it is too sh	Maybe	Maybe	Maybe	Yes	Yes	0	Very coop	Always co	Agree	Yes	Yes	90-100% (a	0
13	3	Female	303:04:00	well enoug	Maybe	Disagree	Maybe	0	Maybe	I agree wit	Maybe	Maybe	Disagree	No	Yes	Choosing b	Very coop	Always co	Some of th	Yes	No	30-60%	0
14	3	Female	303:04:00	well enoug	Maybe	Maybe	Maybe	0	Maybe	I agree wit	Maybe	Maybe	Maybe	Yes	Yes	Study from	Very coop	Very coop	Some of th	Yes	Yes	90-100% (a	Teach us
15	3	Male	03:03.2	well enoug	Agree	Disagree	Agree	0	Agree0	I agree wit	Maybe	Agree	Agree	Yes	No	0	Always co	Always co	Disagree	Yes	No	0-30%	0
16	3	Female	303:04:00	well enoug	Maybe	Disagree	Maybe	Increase e	Agree0	I agree wit	Maybe	Maybe	Agree	Yes	Yes	Reduce	Always co	Very coop	Some of th	Yes	No	30-60%	I think the
17	1	Female	303:04:00	Very well0	Agree	Agree	Agree	0	Agree0	I agree wit	Agree	Agree	Agree	Yes	Yes	0	Always co	Always co	Agree	Yes	Yes	90-100% (a	0
18	1	Female	303:04:00	well enoug	Agree	Agree	Agree	0	Agree0	I agree wit	Agree	Agree	Agree	Yes	Yes	More inter	Always co	Always co	Agree	Yes	Yes	90-100% (a	Problem s
19	1	Female	02:02.9	Very well0	Maybe	Maybe	Maybe	to have	Agree0	I agree wit	Agree	Agree	Agree	Yes	Yes	making the	Very coop	Very coop	Some of th	Yes	Yes	90-100% (a	0
20	1	Female	03:03.2	well enoug	Maybe	Maybe	Maybe	Yes	Maybe	I agree wit	Maybe	Maybe	Maybe	No	No	0	Never coo	Never coo	Some of th	No	No	30-60%	0

Phase 5

Expectation VS Real

Expectation VS Real

Total number of answered surveys = 100

```
]: import pandas as pd
   data = pd.read_csv("C:\\Users\\karas2\\Desktop\\cleandata7.csv")
   data
```

]:

	level	gender	CGPA	theoretical part	material clearness	enough practice examples	material sufficiency	ideas to improve materials	timeframe suitability	CC opnions	...	rooms/Labs are clean	rooms/Labs are big	roon ha
0	2	Female	303:04:00	well enough to pass0	Maybe	Agree	Maybe	It should be more simple and comprehensive	Maybe	It is too long for the semester0	...	No	No	
1	3	Female	303:04:00	Very well0	Maybe	Maybe	Disagree	0	Agree0	I agree with the previous question0	...	No	No	
2	2	Female	303:04:00	well enough to pass0	Agree	Maybe	Maybe	Need to cover more topics or clearer informati...	Agree0	I agree with the previous question0	...	No	No	
3	3	Male	303:04:00	Very well0	Agree	Agree	Disagree	0	Maybe	I agree with the previous question0	...	No	No	
4	3	Female	303:04:00	Very well0	Maybe	Maybe	Maybe	To be more detailed than it is to depend on it	Maybe	I agree with the previous question0	...	Yes	No	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
94	2	Male	303:04:00	well enough to pass0	Agree	Maybe	Disagree	0	Disagree0	It is too long for the semester0	...	No	No	
95	4	Female	03:03.2	well enough to pass0	Maybe	Disagree	Disagree	0	Agree0	I agree with the previous question0	...	No	No	
96	3	Male	02:02.9	I don't understand the theoretical part0	Disagree	Disagree	Disagree	0	Disagree0	It is too long for the semester0	...	No	Yes	
97	2	Male	03:03.2	Very well0	Maybe	Disagree	Maybe	0	Maybe	It is too long for the semester0	...	No	Yes	
98	2	Female	01:01.9	well enough to pass0	Maybe	Maybe	Agree	Books and work shop	Maybe	It is too long for the semester0	...	Yes	Yes	

99 rows × 36 columns

Expectation VS Real

This Code is to compare the answers of 3 pairs of relevant questions to have the knowledge of :

- THE NUMBER OF CARELESS-NESS ANSWERED SURVEYS
- THE NUMBER OF SURVEYS THAT WILL BE TAKEN IN CONSEDERATION

```
In [3]: data['timeframe suitability']=data['timeframe suitability'].replace('Agree0','Yes')
        data['timeframe suitability']=data['timeframe suitability'].replace('Disagree0','No')
        p1=data['timeframe suitability'].str.strip().str.lower()!=data['Semester time  '].str.strip().str.lower()
        p1.value_counts()

Out[3]: True     69
        False    30
        dtype: int64

In [4]: data['Interactive learning  effectiveness']=data['Interactive learning  effectiveness'].replace('Agree','Yes')
        data['Interactive learning  effectiveness']=data['Interactive learning  effectiveness'].replace('Disagree','No')
        p2=data['Interactive learning  effectiveness'].str.strip().str.lower()==data['interactive techniques impact'].str.strip().str.low
        p2.value_counts()

Out[4]: True     69
        False    30
        dtype: int64

In [5]: data['Teamwork ']=data['Teamwork '].replace('Agree','Yes')
        data['Teamwork ']=data['Teamwork '].replace('Disagree','No')
        p3=data['soft skills '].str.strip().str.lower()==data['Teamwork '].str.strip().str.lower()
        p3.value_counts()

Out[5]: True     51
        False    48
        dtype: int64
```

Expectation VS Real

From the previous code we can conclude that :

Number of Carelessness answered Surveys

Equals 30

Number of Surveys that will be taken in consideration

Equals 60

Phase 6

Analyzing data

Descriptive Analysis

We see that all data types are categorical
So we need to deal with it in a different way

Describtive Analysis

In [3]:
```
data.dtypes
```

Out[3]:

level	int64
gender	object
CGPA	object
theoretical part	object
material clearness	object
enough practice examples	object
material sufficiency	object
ideas to improve materials	object
timeframe suitability	object
CC opnions	object
assignments relation to courses	object
assignments measurment of knowledge	object
project discussions efficiency	object
Attending lectures impact	object
Labs/Classes impact	object
suggestions for BL	object
doctors cooporations	object
teaching assistants cooporations	object
interactive teaching techniques	object
interactive techniques impact	object
courses related to jobs	object
courses qualifies students	object
list of IC	object
soft skills	object
soft skills suggistions	object
learning environment can affect performance	object
rooms/Labs are clean	object
rooms/Labs are big	object
rooms/Labs have clear board	object
halls suitable for gaps	object
suggestions for better exp	object

Descriptive Analysis

With frequency we notice that most of students are females which have high CGPA but they still see that materials could be not sufficient enough or not clear enough

Using frequency for describtive analysis

```
In [4]: freq_gender = data['gender'].value_counts()
        freq_gender

Out[4]: Female    53
        Male      46
        Name: gender, dtype: int64

In [5]: freq_CGPA=data['CGPA'].value_counts()
        freq_CGPA

Out[5]: 303:04:00      51
        02:02.9        21
        03:03.2        21
        Less than 1     4
        01:01.9         2
        Name: CGPA, dtype: int64

In [6]: freq_material_sufficiency =data['material sufficiency '].value_counts()
        freq_material_sufficiency

Out[6]: Maybe       50
        Disagree    33
        Agree       16
        Name: material sufficiency , dtype: int64

In [7]: freq_material_clearness=data['material clearness'].value_counts()
        freq_material_clearness

Out[7]: Maybe       60
        Agree       27
        Disagree    12
        Name: material clearness, dtype: int64
```

Descriptive Analysis

We also notice that there might be not enough practical examples although the time frame is quiet suitable so professors must consider increasing the practical examples

```
In [8]: freq_enough_practice_examples=data['enough  practice examples'].value_counts()
        freq_enough_practice_examples

Out[8]: Maybe       44
        Disagree    30
        Agree       25
        Name: enough  practice examples, dtype: int64

In [9]: freq_timeframe_suitability=data['timeframe suitability'].value_counts()
        freq_timeframe_suitability

Out[9]: Agree0       40
        Maybe        39
        Disagree0    20
        Name: timeframe suitability, dtype: int64
```

Descriptive Analysis

With mode we noticed that assignments are related to courses but they are not enough to measure the knowledge of students

Using Mode for Descriptive Analysis

```
In [10]: mode_assignments_relation_to_courses  = data['assignments relation to courses'].mode()
         mode_assignments_relation_to_courses

Out[10]: 0    Agree
         dtype: object

In [11]: mode_assignments_measurment_of_knowledge=data['assignments measurment of knowledge'].mode()
         mode_assignments_measurment_of_knowledge

Out[11]: 0    Maybe
         dtype: object
```

Descriptive Analysis

we also noticed that students see that working environment can affect their performance and they think that halls are clean and have clear boards but they are not big enough or suitable for gaps

```
In [12]: mode_learning_environment_can_affect_performance=data['learning environment can affect  performance'].mode()
         mode_learning_environment_can_affect_performance

Out[12]: 0    Yes
         dtype: object

In [13]: mode_rooms_Labs_are_clean=data['rooms/Labs are clean '].mode()
         mode_rooms_Labs_are_clean

Out[13]: 0    Yes
         dtype: object

In [14]: mode_rooms_Labs_are_big=data['rooms/Labs are big  '].mode()
         mode_rooms_Labs_are_big

Out[14]: 0    No
         dtype: object

In [15]: mode_rooms_Labs_have_clear_board =data['rooms/Labs have clear board'].mode()
         mode_rooms_Labs_have_clear_board

Out[15]: 0    Yes
         dtype: object

In [16]: mode_halls_suitable_for_gaps =data['halls suitable for gaps '].mode()
         mode_halls_suitable_for_gaps

Out[16]: 0    No
         dtype: object
```

Descriptive Analysis

Using Percentages for Descriptive Analysis

In [17]:
```
percentages_soft_skills  = data['soft skills '].value_counts(normalize=True) * 100
percentages_soft_skills
```

Out[17]:
```
Yes    52.525253
No     47.474747
Name: soft skills , dtype: float64
```

we notice that almost 52% of students have soft skills but there are alot of other students need to build thier soft skills

In [18]:
```
percentages_courses_related_to_jobs=data['courses related to jobs'].value_counts(normalize=True) * 100
percentages_courses_related_to_jobs
```

Out[18]:
```
No     59.59596
Yes    40.40404
Name: courses related to jobs, dtype: float64
```

we notice that almost 60% of students see that courses are not related to jobs

Descriptive Analysis

Using Cross Tabulation (Contingency Table) for Descriptive Analysis

In [19]:
```
cross_gen_gpa = pd.crosstab(data['gender'], data['CGPA'])
cross_gen_gpa
```

Out[19]:

gender	01:01.9	02:02.9	03:03.2	303:04:00	Less than 1
Female	1	8	14	28	2
Male	1	13	7	23	2

we notice that females have higher CGPA than males so may be gender can affect CGPA

In [20]:
```
cross_doc_ass_coop=pd.crosstab(data['doctors cooporations '],data['teaching assistants cooporations'])
cross_doc_ass_coop
```

Out[20]:

doctors cooporations	Always cooperative0	Never cooperative0	Very cooperative in campus, not so cooperative online0	Very cooperative online, not so cooperative in campus0
Always cooperative0	22	2	4	0
Never cooperative0	2	5	5	2
Very cooperative in campus, not so cooperative online0	22	2	17	7
Very cooperative online, not so cooperative in campus0	1	0	3	5

we notice the relation between doctors and assistants cooperation in helping students which mean that in most cases a cooporative doctor has a cooperative assistant

Descriptive Analysis

In [21]:
```
cross_int_tech_impact=pd.crosstab(data['interactive teaching techniques'],data['interactive techniques impact'])
cross_int_tech_impact
```

Out[21]:

interactive techniques impact	No	Yes
interactive teaching techniques		
Agree	2	15
Disagree	2	12
Some of them	4	64

we notice that interactive teaching techniques have high impact on students

Inferential Statistics

Inferential Statistics

Chi_square Test for Inferential Statistics

In [22]:
```
import scipy.stats as stats
```

In [23]:
```
#Perform a chi-square test to determine if there is a significant association between the two categorical variables.
#The chi-square test evaluates the null hypothesis that the two variables are independent.
chi2, p_value,gen,gpa = stats.chi2_contingency(cross_gen_gpa)
print("Chi-square value:", chi2)
print("p-value:", p_value)
```

```
Chi-square value: 3.5367379929822627
p-value: 0.47231428372065243
```

we will assume that significance level=0.05 so from the above p_value, we will reject null hypothesis and accept the alternative one which prove that there is a relation between gender and gpa

In [24]:
```
chi2, p_value,int_,tech = stats.chi2_contingency(cross_int_tech_impact)
print("Chi-square value:", chi2)
print("p-value:", p_value)
```

```
Chi-square value: 1.4787376489057165
p-value: 0.4774151532146448
```

we will assume that significance level=0.05 so from the above p_value, we will reject null hypothesis and accept the alternative one which prove that interactive teaching techniques have impact on students

Correlation Analysis

To measure the correlation coefficient between two categorical variables, you can use a statistical measure called Cramer's V. Cramer's V is a widely used measure of association between two categorical variables. It ranges from 0 to 1, with higher values indicating stronger association between the variables.***

The main assumptions are:

(1)The observations are independent.

(2)The expected frequency count for each cell in the contingency table is at least 5

Correlation Analysis

Conclusion: No Correlation Between Level & Material Clearness. Therefore, There is no Change in Material Perception as The Years Passes by

In [25]:
```
from scipy.stats import chi2_contingency
import numpy as np
```

In [26]:
```
# create a contingency table of the level and material clearness variables
table = pd.crosstab(data['level'], data['material clearness'])

# calculate the chi-squared statistic and related values
chi2, p, dof, expected = chi2_contingency(table)

# calculate Cramer's V
n = np.sum(table)
phi2 = chi2/n
r, k = table.shape
phi2corr = np.maximum(0, phi2 - ((k-1)*(r-1))/(n-1))
r_corr = r - ((r-1)**2)/(n-1)
k_corr = k - ((k-1)**2)/(n-1)
v = np.sqrt(phi2corr/np.minimum((k_corr-1),(r_corr-1)))

# print the contingency table and Cramer's V
print(table)
print("Cramer's V:", v)
```

```
material clearness  Agree  Disagree  Maybe
level
1                       8         1     15
2                       8         4     15
3                      10         4     25
4                       1         3      5
Cramer's V: material clearness
Agree       0.056501
Disagree    0.000000
Maybe       0.049859
dtype: float64
```

Correlation Analysis

Conclusion: High Level Correlation Between High Satisfaction And Level

In [27]:
```
# create a contingency table of the level and overall satisfaction variables
table = pd.crosstab(data['level'], data['overall satisfaction  '])

# calculate the chi-squared statistic and related values
chi2, p, dof, expected = chi2_contingency(table)

# calculate Cramer's V
n = np.sum(table)
phi2 = chi2/n
r, k = table.shape
phi2corr = np.maximum(0, phi2 - ((k-1)*(r-1))/(n-1))
r_corr = r - ((r-1)**2)/(n-1)
k_corr = k - ((k-1)**2)/(n-1)
v = np.sqrt(phi2corr/np.minimum((k_corr-1),(r_corr-1)))

# print the contingency table and Cramer's V
print(table)
print("Cramer's V:", v)
```

```
overall satisfaction    0-30%  30-60%  60-90%  90-100%
level
1                           2       9      12        1
2                           4       7      15        1
3                           7      18       9        5
4                           4       4       1        0
Cramer's V: overall satisfaction
0-30%      0.431343
30-60%     0.276847
60-90%     0.280789
90-100%    0.803103
dtype: float64
```

Regression Analysis

linear Regression

In [28]:
```
# Convert the 'material clearness' column to binary variables
data[['material_clearness_Maybe', 'material_clearness_Agree', 'material_clearness_Disagree']] = pd.get_dummies(data['material cle

# Convert the 'CGPA' column to numerical values
cgpa_categories = ['Less than 1', '1:00:00-1:59:59', '2:00:00-2:59:59', '3:00:00-3:59:59', '4:00:00-4:59:59', '5:00:00-5:59:59']
data['CGPA'] = pd.Categorical(data['CGPA'], categories=cgpa_categories)
data['CGPA'] = data['CGPA'].cat.codes

# Define the variables
x = data[['material_clearness_Maybe', 'material_clearness_Agree', 'material_clearness_Disagree']]
y = data['CGPA']

# Perform the regression analysis
from sklearn.linear_model import LinearRegression
model = LinearRegression()
model.fit(x, y)

# Print the coefficients
print('Intercept:', model.intercept_)
print('Coefficients:', model.coef_)

# Make a prediction
prediction = model.predict(pd.DataFrame({'material_clearness_Maybe': [0], 'material_clearness_Agree': [1], 'material_clearness_Di
print('Prediction:', prediction)
```

```
Intercept: 8401039014246.986
Coefficients: [-8.40103901e+12 -8.40103901e+12 -8.40103901e+12]
Prediction: [-0.75]
```

Regression Analysis

The intercept is the value of the dependent variable (y) when all of the independent variables (x) are equal to 0.
it is the value of y that we would expect when all of the independent variables have no effect on the dependent variable.
The intercept is also sometimes called the constant term or the bias term.

The coefficients represent the change in the dependent variable (y) for a one-unit increase in each
of the independent variables (x).
predictions for new observations by estimating the value of the dependent variable based on the values independent variables.
this indicates that the linear regression model doesn't fit so we will use the logistic model

Regression Analysis

In [32]:
```
# Load the data
data = pd.read_csv("C:\\Users\\Kimo Store\\Downloads\\Telegram Desktop\\cleandata7.csv")

# Convert the 'material clearness' column to binary variables
data[['material_clearness_Maybe', 'material_clearness_Agree', 'material_clearness_Disagree']] = pd.get_dummies(data['material cle

# Convert the 'CGPA' column to numerical values
cgpa_categories = ['Less than 1', '1:00:00-1:59:59', '2:00:00-2:59:59', '3:00:00-3:59:59', '4:00:00-4:59:59', '5:00:00-5:59:59']
data['CGPA_cat'] = pd.Categorical(data['CGPA'], categories=cgpa_categories).codes

# Define the variables
x = data[['material_clearness_Maybe', 'material_clearness_Agree', 'material_clearness_Disagree']]
y = data['CGPA_cat']

# Split data into training and testing sets
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=42)

# Fit a logistic regression model
model = LogisticRegression(random_state=42)
model.fit(x_train, y_train)

# Print the intercept and coefficients for each category of the dependent variable
print('Intercepts:', model.intercept_)
print('Coefficients:', model.coef_)

# Make a prediction for a new observation
new_observation = pd.DataFrame({'material_clearness_Maybe': [0], 'material_clearness_Agree': [1], 'material_clearness_Disagree':
new_obs_pred = model.predict(new_observation)
#print('New observation:', new_observation.to_dict())
print('Prediction:', new_obs_pred)
```

```
Intercepts: [-2.79527803]
Coefficients: [[-0.7039678   1.24684441 -0.54287659]]
Prediction: [-1]
```

Data Validation

We noticed that accuracy didn’t give us an accurate answer so we will use another method called cross validation score

Validation

In [49]:
```
from sklearn.metrics import accuracy_score, precision_score, recall_score

y_pred = model.predict(x_test)

accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred)
recall = recall_score(y_test, y_pred)
f1 = f1_score(y_test,y_pred)

print("Accuracy:", accuracy)
print("Precision:", precision)
print("Recall:", recall)
```

```
Accuracy: 1.0
Precision: 0.0
Recall: 0.0
```

Data Validation

It gave us a better result

```
In [34]: from sklearn.model_selection import cross_val_score

         scores = cross_val_score(model, x, y, cv=5)
         print("Cross-Validation Scores:", scores)
         print("Mean CV Score:", scores.mean())

         Cross-Validation Scores: [0.95 0.95 0.95 0.95 1.  ]
         Mean CV Score: 0.96
```

Data Reliability

We notice that reliability score is relatively low so we will drop some columns and try to improve it

Calculating Reliability

In [72]:
```
!pip install krippendorff
from krippendorff import alpha

alpha_value = alpha(data_encode)
alpha_value
```

```
Requirement already satisfied: krippendorff in c:\program files\anaconda\lib\site-packages (0.6.0)
Requirement already satisfied: numpy<2.0,>=1.21 in c:\program files\anaconda\lib\site-packages (from krippendorff) (1.24.3)
```

Out[72]: 0.340487197998403

Data Reliability

We see that it's much better now

```
In [80]: data_encode1=data_encode.drop(['material clearness','enough  practice examples','material sufficiency ','timeframe suitability','
         data_encode1
```

Out[80]:

	gender	CGPA	theoretical part	courses related to jobs	rooms/Labs are big	halls suitable for gaps
0	0	3	3	1	0	0
1	0	3	2	0	0	0
2	0	3	3	1	0	0
3	1	3	2	1	0	0
4	0	3	2	0	0	0
...	...	...	...	...	...	...
94	1	3	3	0	0	1
95	0	2	3	0	0	0
96	1	1	1	0	1	0
97	1	2	2	1	1	1
98	0	0	3	0	1	0

99 rows × 6 columns

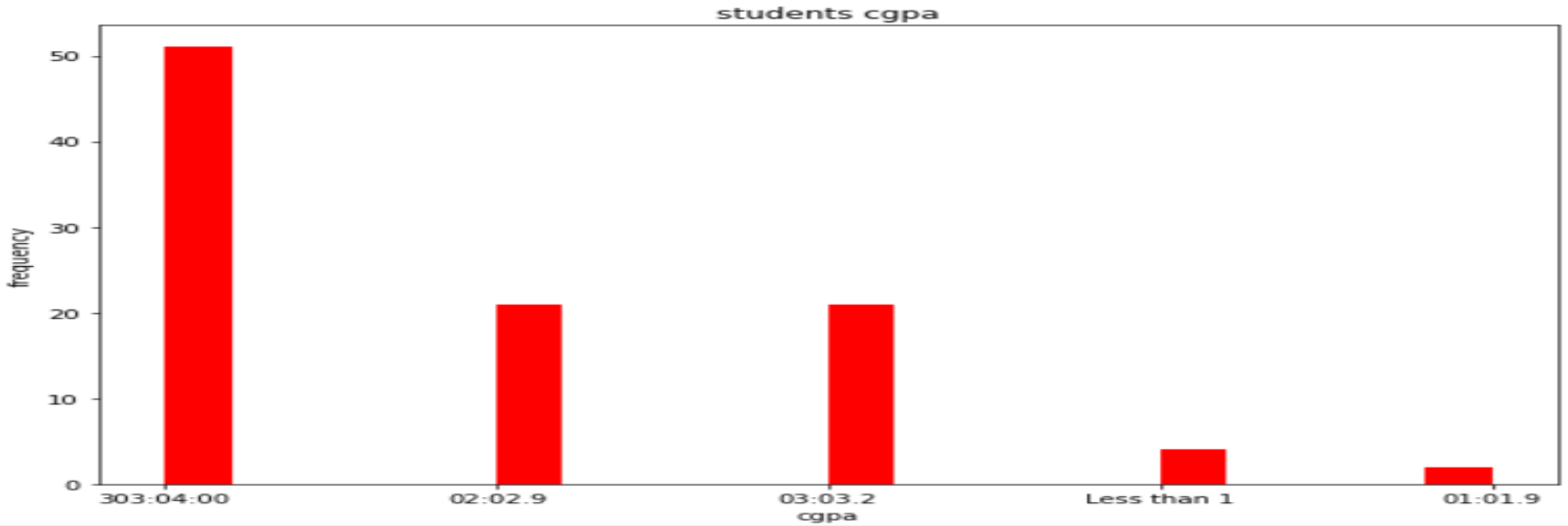
```
In [81]: alpha_value = alpha(data_encode1)
         alpha_value
```

```
Out[81]: 0.7268448260842052
```

Data Visualization

```
plt.figure(figsize=(10,7))
x = data["CGPA"]
plt.hist(x, bins= 20, color="red")
plt.title("students cgpa")
plt.xlabel("cgpa")
plt.ylabel("frequency")
plt.show
```

```
|: <function matplotlib.pyplot.show(close=None, block=None)>
```

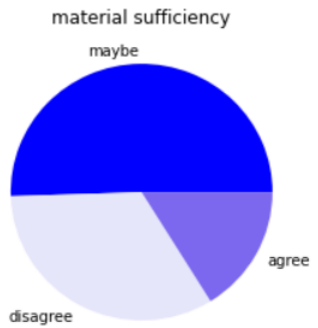


Data Visualization

In the last two plots above we can that the majority of students tend to think that the courses materials are not very great also see that it does not meet jobs needs so, one of the things that can be taken in consideration is updating the courses content

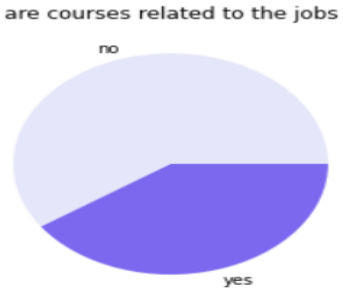
```
In [41]: plt.pie([50,33,16], labels=['maybe','disagree','agree'], colors=['#0000FF','#E6E6FA','#7B68EE'])
         plt.title('material sufficiency')
         plt.show
```

```
Out[41]: <function matplotlib.pyplot.show(close=None, block=None)>
```




```
In [42]: plt.pie([59,40], labels=['no','yes'], colors=['#E6E6FA','#7B68EE'])
         plt.title('are courses related to the jobs')
         plt.show
```

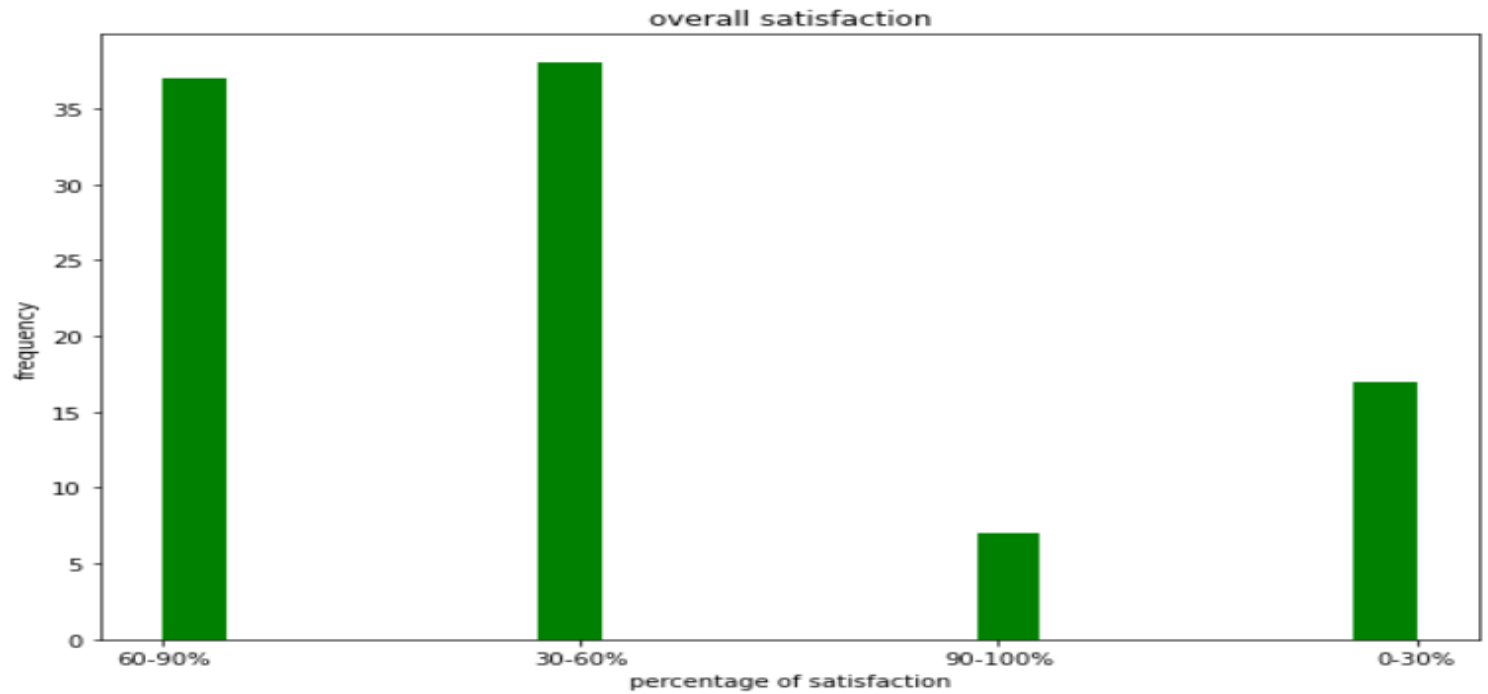
```
Out[42]: <function matplotlib.pyplot.show(close=None, block=None)>
```



Data Visualization

```
In [43]: plt.figure(figsize=(10,7))
         x = data["overall satisfaction  "]
         plt.hist(x, bins= 20, color="green")
         plt.title("overall satisfaction")
         plt.xlabel("percentage of satisfaction")
         plt.ylabel("frequency")
         plt.show
```

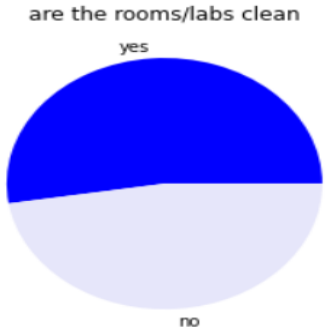
```
Out[43]: <function matplotlib.pyplot.show(close=None, block=None)>
```



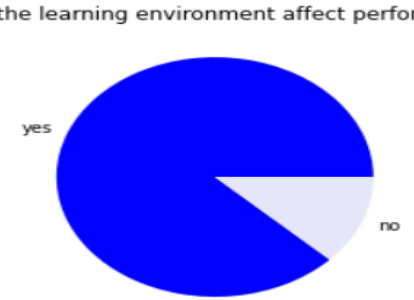
Data Visualization

In the last 3 plots above we see that nearly all the students agree that the learning environment does affect the performance and they do not see that the rooms or the labs are big enough and half of them see that it is not even clean so, working on the rooms and the labs maybe able to help improve students performance

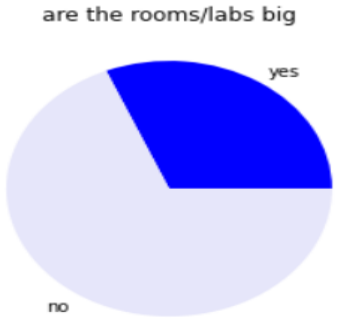
```
In [46]: plt.pie([52,47], labels=['yes','no'], colors=['#0000FF','#E6E6FA'])
         plt.title('are the rooms/labs clean')
         plt.show
Out[46]: <function matplotlib.pyplot.show(close=None, block=None)>
```




```
In [44]: plt.pie([87,12], labels=['yes','no'], colors=['#0000FF','#E6E6FA'])
         plt.title('does the learning environment affect performance')
         plt.show
Out[44]: <function matplotlib.pyplot.show(close=None, block=None)>
```



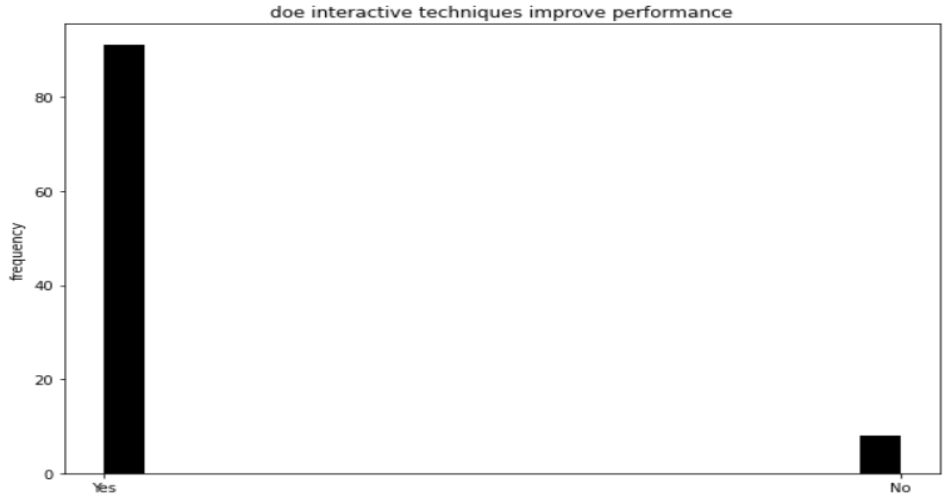

```
In [45]: plt.pie([31,68], labels=['yes','no'], colors=['#0000FF','#E6E6FA'])
         plt.title('are the rooms/labs big')
         plt.show
Out[45]: <function matplotlib.pyplot.show(close=None, block=None)>
```



Data Visualization

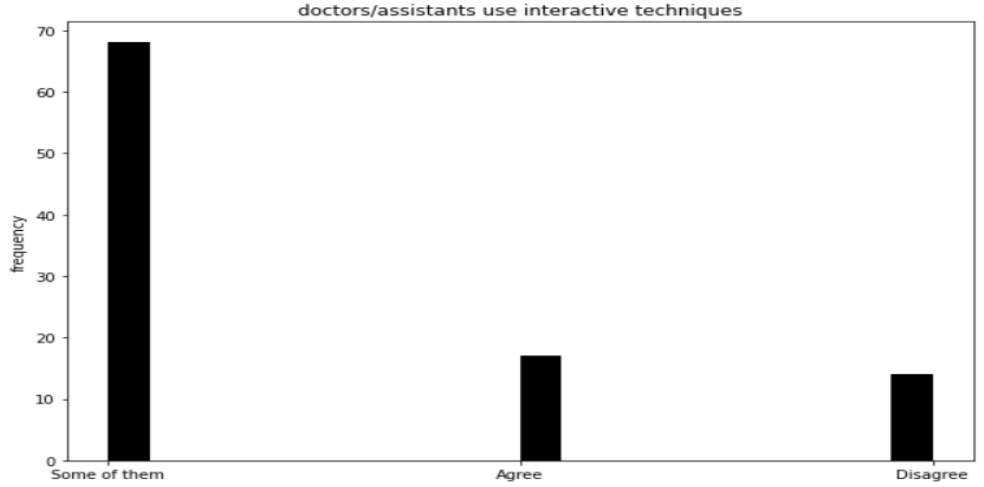
```
In [48]: plt.figure(figsize=(10,7))
         x = data["interactive techniques impact"]
         plt.hist(x, bins= 20, color="black")
         plt.title("doe interactive techniques improve performance")
         plt.xlabel("")
         plt.ylabel("frequency")
         plt.show
```

```
Out[48]: <function matplotlib.pyplot.show(close=None, block=None)>
```




```
In [47]: plt.figure(figsize=(10,7))
         x = data["interactive teaching techniques"]
         plt.hist(x, bins= 20, color="black")
         plt.title("doctors/assistants use interactive techniques")
         plt.xlabel("")
         plt.ylabel("frequency")
         plt.show
```

```
Out[47]: <function matplotlib.pyplot.show(close=None, block=None)>
```



In the above two plots we see that nearly all the students can agree that using interactive techniques can improve performance and see that not all of the doctors or assistants use these techniques

Summary

from the above we concluded that :

1) gender might affect CGPA , so we need to search deeper and find the reason

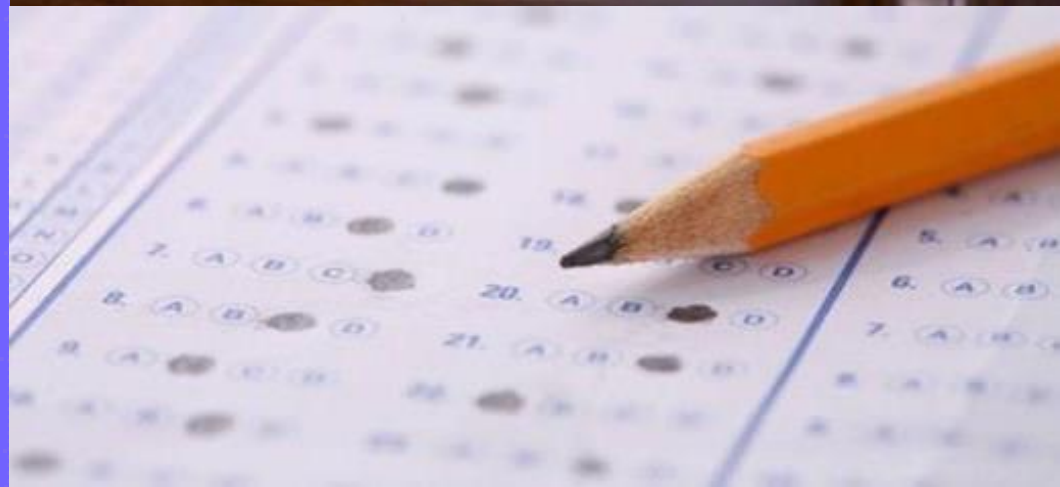
2) Interactive teaching techniques are quite important so we need to focus on changing the traditional techniques

3) Halls are clean but they need to be bigger and we need to increase suitable places for gaps

4) Cooperative professor often means cooperative staff, so we must focus on choosing the right professor and staff

5) Semester time is good but practical examples are not enough so we might consider increasing practical examples

6) Increasing practical examples might increase material clearness as it's not very good



Thank you