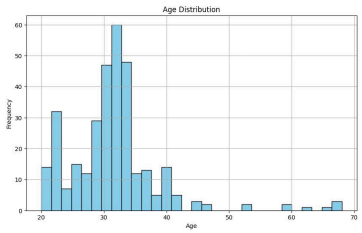
Data Analysis Assignment
Group M

Team Members:

1. M.A.S.N Aththnayake - 210055J
2. Nayanathara P.M.C. - 210417X
3. Janeesha Wickramasinghe - 210706H
4. M.A.C.L Mallikarachchi - 210363C
5. Pranavan Subendiran - 210491P

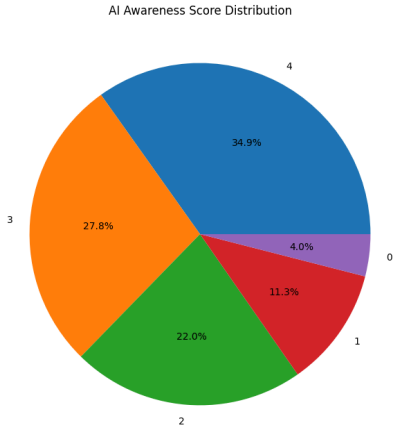
Descriptive analysis

Gender	Male: 53% Female: 47%
Age	Mean age: 31 
Educational Level	More than 75% have completed undergraduate and postgraduate degrees
Field of Occupation	Majority are working in Architecture, Engineering, Computer and Mathematical fields
Occupation level	Middle level is the most prominent
Country of residence	Majority of participants are from Sri Lanka (200+ out of 327)

Insights gained

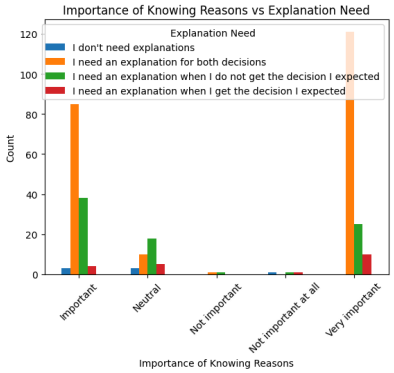
1. **Individual awareness on AI**
   Section 2 of the survey assessed the AI awareness of the participants.
   Here, in order to understand how far the survey participants are aware of the concept of Artificial Intelligence we took the relevant section out, encoded the correct answers with 1 (a positive value) and the wrong answers with 0.

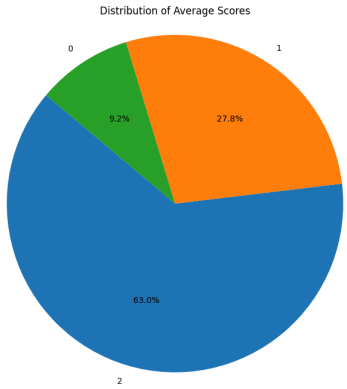
Finally we calculated the sum of all the scores of the questions and obtained a visual representation (a pie chart).
4/4, 3/4 , 2/4, 1/4 and 0 based on the scores out of 4 questions



2. Interest of participants towards XAI

The portion in blue (63 % of the total population) depicts the total population that consider reasoning behind AI important or very important and also require knowing the reason for both when they get the expected decision and when they don't get the expected decision.

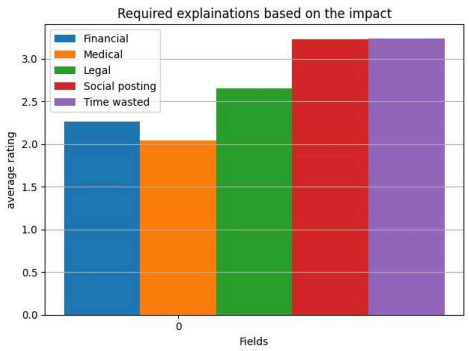




3. Calculated mean of the importance of XAI explanations in different fields

The values were taken 1-most TO 5-least and an average was taken for all participant answers for question 17 on the impact of XAI with regards to different fields.
E.g: Financial, Medical, Legal, Social posting, Time wasted

We can clearly see that survey participants are more inclined towards knowing and understanding the impacts made by Financial and Medical sections.



Hypothesis testing

Tools and Techniques:

- Significance level- 5%
- Tests
  - Chi Square test of independence (Categorical- Categorical)
  - F test /ANOVA test (Categorical- Numerical)
- Python libraries used- Pandas, Scipy

AI Awareness Score:

We calculated it based on questions 7-10 by allocating positive scores for correct answers. (function given below)

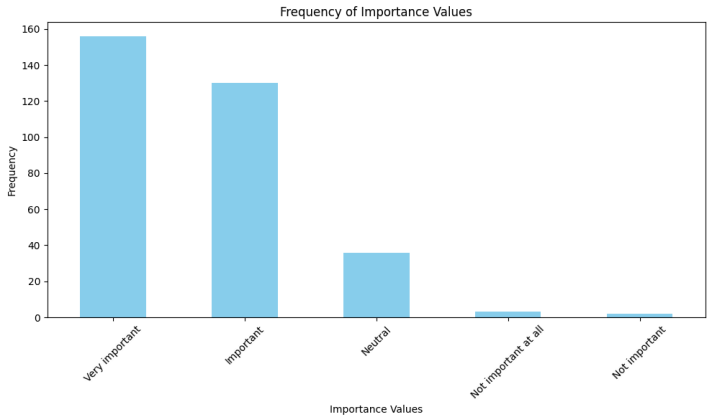
```
def calc_ai_awareness_score(df):
  for index, row in df.iterrows():
    score = 0
    if (row['ai_in_app'] == 'Email'): score += 1
    if (row['ai_pneumonia_detection'] == 'Yes'): score += 1
    if (row['ai_pin_unlock'] == 'No'): score += 1
    if (row['ai_ads'] == 'Yes'): score += 1

    df.loc[index, 'ai_awareness_score'] = score
  df.drop(columns=['ai_in_app', 'ai_pneumonia_detection', 'ai_pin_unlock', 'ai_ads'], axis=1, inplace=True)
  return df
```

Hypothesis Testing:

Tests	Hypothesis	Conclusions
1. **Gender significantly impacts the perceived importance of XAI.**	H0: Gender doesn't impact the perceived importance of XAI.(independency) Ha: Gender significantly impacts the perceived importance of XAI.(dependency)	Chi-Squared Test Results: Chi-Squared Statistic: 7.299505680817227 p-value: 0.5046898643051009 Degrees of Freedom: 8 Hypothesis testing result: Failed to reject H0 at 5.0% significance level
2. **Age significantly impacts the perceived importance of XAI**	H0: Age doesn't impact the perceived importance of XAI.(independency) Ha: Age significantly impacts the perceived importance of XAI.(dependency)	F Test Results: F-statistic: 1.1944974862665445 p-value: 0.31313028666404996 Hypothesis testing result: Failed to reject H0 at 5.0% significance level
3. **Education significantly impacts the perceived importance of XAI**	H0: Education level doesn't impact the perceived importance of XAI.(independency) Ha: Education level significantly impacts the perceived importance of XAI.(dependency)	Chi-Squared Test Results: Chi-Squared Statistic: 16.43986508582492 p-value: 0.6889882655515351 Degrees of Freedom: 20 Hypothesis testing result: Failed to reject H0 at 5.0% significance level
4. **Occupation significantly impacts the perceived importance of XAI**	H0: Field of occupation doesn't impact the perceived importance of XAI.(independency) Ha: Field of occupation significantly impacts the perceived importance of XAI.(dependency)	Chi-Squared Test Results: Chi-Squared Statistic: 46.38220459924884 p-value: 0.9918259275396065 Degrees of Freedom: 72 Hypothesis testing result: Failed to reject H0 at 5.0% significance level
5. **Career level significantly impacts the perceived importance of XA**I	H0: Career level doesn't impact the perceived importance of XAI.(independency) Ha: Career level significantly impacts the perceived importance of XAI.(dependency)	Chi-Squared Test Results: Chi-Squared Statistic: 47.035688083592 p-value: 0.34930628886829973 Degrees of Freedom: 44 Hypothesis testing result: Failed to reject H0 at 5.0% significance level
6. **AI awareness significantly impacts the perceived importance of XAI**	H0: AI awareness doesn't impact the perceived importance of XAI.(independency) Ha: AI awareness significantly impacts the perceived importance of XAI.(dependency)	F Test Results: F-statistic: 1.9304332356129779 p-value: 0.10515075252741682 Hypothesis testing result: Failed to reject H0 at 5.0% significance level

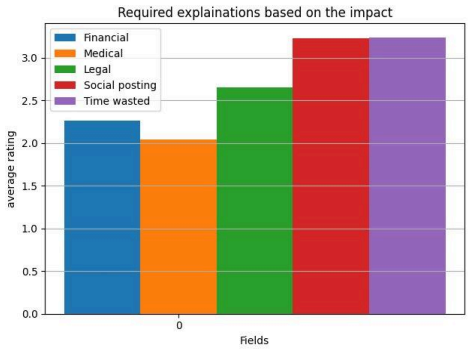
I. Is it important for users to know the reasons behind the decisions of AI systems?



Question 11 discusses the importance of XAI: Reasons behind decisions made by AI, for users.

It is evident from this bar graph that it is important or very important for a majority to know explanations.

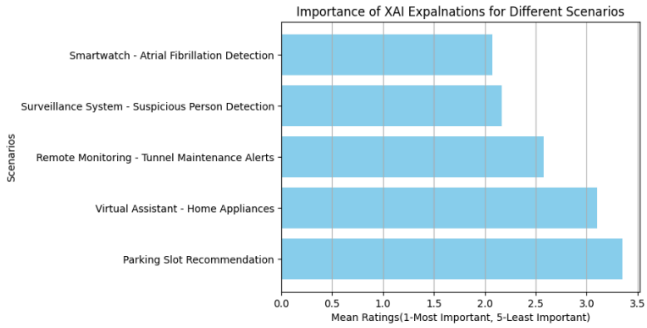
II. Which impact (financial, medical, legal, etc.) caused by AI decisions gives rise to XAI being of higher importance?



Calculated mean of the importance of XAI explanations in different fields
The values were taken 1-most TO 5-least and an average was taken for all participant answers for question 17 on the impact of XAI with regards to different fields.
E.g: Financial, Medical, Legal, Social posting, Time wasted

We can clearly see that survey participants are more inclined towards knowing and understanding the impacts made by Financial and Medical sections.

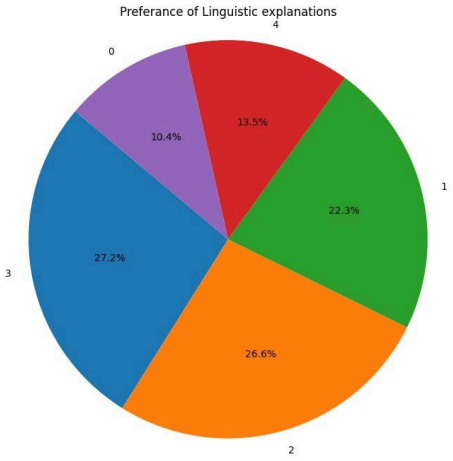
III. Which ELE application gives rise to XAI being of higher importance?



Calculated mean of the importance of XAI explanations for the mentioned scenarios.
The values were taken 1-most TO 5-least and an average was taken for all responses of participants for question 20 on the importance of XAI for each scenario.
E.g: Smartwatch, Surveillance system, Remote monitoring, Virtual assistant, Parking slot recommendation.

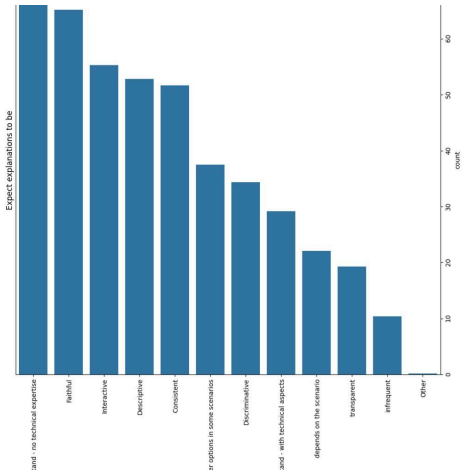
Overall, the data suggests that scenarios involving healthcare, public safety, and critical infrastructure maintenance are perceived as applications that give rise to XAI being of higher importance compared to scenarios involving everyday tasks or conveniences.

IV.Are linguistic explanations preferred over other types of explanations?



YES.
In the questions(Q13- Q16) we can clearly see that the linguistic explanation is given by answer (b) whereas a rather technical explanation is given by answer (a). The pie chart depicts the number of times (b) has been selected as the answer by users. (e.g: selected (b) for all 4 [RED], selected (b) for only 3 [BLUE] etc…)
And as seen more than 50% have selected answer (b) for two or more questions out of the four.

V. RQ4: What are the expected characteristics of explanations?



Question 21 asks the participants to specify the expected characteristic of explanations and when plotted in a bar chart, it's clear that **being easy to understand**, **faithfulness**, **interactivity**, **descriptiveness** and **consistency** were required by a majority.

Google Colab Links

1. Analysis:
   https://colab.research.google.com/drive/12PboGrfMLOV5vmjS9rlxhJG_kdsq6W7C?usp=sharing

2. Hypothesis Testing:
   https://colab.research.google.com/drive/1MsBjJefUWBwT2MDoXCsKNNo_hh_ueHXj?usp=sharing

References

https://youtube.com/playlist?list=PLiC1doDIe9rCYWmH9wIEYEXXaJ4KAi3jc&si=ZXERG-U_30t1plqa
https://www.geeksforgeeks.org/how-to-perform-a-one-way-anova-in-python/
https://www.pythonfordatascience.org/chi-square-test-of-independence-python/