PROCEDURE

KEIGO DEFINITION AND COUNT

- I first define some commonly occurring keigos and write the patterns under each category (teineigo, sonkeigo, kenjougo) so that I can find these patterns in the words.

```python
# Teineigo patterns
teineigo_patterns = [
    r"\bo-?\w+",                              # Words starting with "o-" (e.g., o-hanashi)
    r"\bgo-?\w+",                             # Words starting with "go-" (e.g., go-zonji)
    r"\b\w+ mas(u|u)?\b",                     # Words ending with "masu" or "mas" (e.g., tabemasu)
    r"\b\w+ desu\b",                          # Words ending with "desu" (e.g., sensei desu)
    r"\b\w+ de gozaimasu\b",                  # Formal variation of "desu"
    r"\b\w+ gozaimasu\b",                     # Words ending with "gozaimasu"
]


sonkeigo_patterns = [
    r"\bitadak[ui]\b",  # Capture standalone "itadaku"
    r"\bsonchō shimasu\b",  # Capture "sonchō shimasu"
    r"\bonegai mōshiagemasu\b",  # Capture "onegai mōshiagemasu"
    r"\w+ nasu\b",  # Words ending with "nasu"
    r"\w+ kudasai\b",  # Words ending with "kudasai"
    r"\w+ irasshaimasu\b",  # Capture "irasshaimasu" specifically
    r"\w+ ni naru\b",  # Words ending with "ni naru"
    r"\w+ itasu\b(?=(?!\bmai\s?|\bkenjōgo\s?)\w+)",  # Words ending with "itasu" (excluding Kenjougo and Mai patterns)
]

kenjougo_patterns = [
    r"\bmai\s?\w+",  # Words starting with "mai"
    r"\w+ suru\b",  # Words ending with "suru"S
    r"\w+ shimasu\b",  # Common action verb in Kenjougo
    r"\w+ itasu\b(?=(?!\bsonkeigo\s?)\w+)",  # Words ending with "itasu" (excluding Sonkeigo patterns)
    r"\w+ ageru\b",  # Words ending with "ageru"
    r"\w+ moshi ageru\b",  # Words ending with "moshi ageru"
    r"\w+ zonjimasu\b",  # Words ending with "zonjimasu"
    r"\w+ ukagaimasu\b",  # Irregular Kenjougo form
]
```

```python
# Irregular patterns
irregular_teineigo_patterns = []  # Empty as per your specification

irregular_sonkeigo_patterns = [
    r"\w+ rareru\b",                          # Irregular Sonkeigo form (e.g., o kake rareru)
    r"\w+ iru\b",                             # Irregular Sonkeigo form (e.g., o ide iru)
    r"\w+ ni naru\b",                         # Irregular Sonkeigo form (e.g., o kai ni naru)
    r"\w+ sha\b",                             # Irregular Sonkeigo form for company names (e.g., gakusha)
]

irregular_kenjougo_patterns = [
    r"\w+ mairu\b",                           # Irregular Kenjougo form (e.g., mairimasu)
    r"\w+ ukagau\b",                          # Irregular Kenjougo form (e.g., o ukagai suru)
    r"\w+ jimasu\b",                          # Irregular Kenjougo form (e.g., zonjimasu)
    r"\w+ sha\b",                             # Irregular Kenjougo form for company names (e.g., heisha)
    r"\w+ meshiagaru\b",                      # Irregular Kenjougo form (e.g., o meshiagari kudasai)
    r"\w+ haidoku\b",                         # Irregular Kenjougo form (e.g., o haidoku suru)
]
```

I tried to make the list as accurate and exhaustive as possible.

- Next, I create a function to **count** the occurrences of each type of keigo and store the counts in three variables: "**teineigo**", "**sonkeigo**", and "**kenjougo**".
- In the function, I convert the sentence to lowercase for easier matching.
- Then, I use **regex** to search for patterns that match each keigo type in the sentence.
- Finally, I return the counts for each type of keigo.

KEIGO CLASSIFICATION

- I created a function to classify the level of keigo (polite language) usage based on the counts of different keigo types.
- First, I set a default message saying "No keigo detected" and create a list to store feedback messages.
- Then, I check if each type of keigo (teineigo, sonkeigo, kenjougo) is used:
  - If **teineigo** is used, I add a message indicating **politeness**; if not, I note that no polite language was detected.
  - If **sonkeigo** is used, I add a message about showing **respect**; otherwise, I mention that respectful language wasn't detected.
  - If **kenjougo** is used, I add a message about **humility**; if not, I indicate that humble language was not found.
- Based on the combination of keigo types detected, I assign a proficiency level:
  - If only **one** type is used, I classify the usage as "**Basic**."
  - If **two** types are used, I classify the usage as "**Intermediate**."
  - If all **three** types are used, I classify it as "**Advanced**."
  - If **none** are detected, I confirm "**No keigo detected**."
- Finally, I combine all the messages into one feedback and return both the proficiency level and the feedback message.

KEIGO OUTPUT

- I translate the input Japanese sentence into Romaji (phonetic representation of Japanese using the Latin alphabet).
- Then, I use the `count_occurrences` function to detect how many keigo forms are present in the Romaji sentence.
- I classify the keigo usage based on the counts, assigning a proficiency level and generating feedback.
- Next, I create a detailed prompt for a language model (LLM) to generate a JSON response. This response contains information about the keigo detected, its contextual validity, a keigo analysis, and tips for improvement.

The LLM is prompted to follow a specific format and generate a JSON response that includes:

- Whether the keigo usage is contextually valid.
- An analysis of whether the keigo usage is appropriate.
- The count of keigo forms detected.
- The feedback and proficiency level generated from the previous functions.
- Improvement tips for better keigo usage.

- Finally, I generate the output in JSON format.

KEIGO CLASSIFICATION EXAMPLES

SAMPLE INPUT 1

```
#n1 advance

sentence1 = """
私は、問題解決力とコミュニケーション能力を活かせると考えています。学生時代にチームで課題
に取り組む際、問題点を見つけて解決策を提案することで、チーム全体をまとめることができまし
た。
"""
```

OUTPUT

Romaji: Watashi wa, mondaikaiketsu-ryoku to komyunikēshon nōryoku o ikaseru to kangaete imasu. Gakusei jidai ni chīmu de kadai ni torikumu sai, mondaiten o mitsukete kaiketsusaku o teian suru koto de, chīmu zentai o matomeru koto ga dekimashita.

```
{
  "contextual_validity_of_keigo": "No",
  "contextual_analysis_of_keigo": "While you used kenjougo, which is
generally appropriate for expressing humility, it might not be the most
suitable in this introductory context. In a self-introduction,  teineigo
(polite language) is usually expected to show respect to the listener. ",
  "keigo_count": {
    "teineigo": 0,
    "sonkeigo": 0,
    "kenjougo": 1
  },
  "keigo_analysis": "No teineigo (polite language) detected. No sonkeigo
(respectful language) detected. You used kenjougo, which reflects humility
in your speech. This is great for situations where you need to humble
yourself. Your kenjougo usage is humble, but adding polite and respectful
forms will make your communication more balanced.",
  "proficiency_level": "Basic (Kenjougo)",
  "improvement_tips": [
    "You should use more teineigo (polite language) when introducing
yourself, especially to someone you don't know well.",
    "Consider using teineigo in future introductions to show courtesy and
respect."
  ]
}
```

SAMPLE INPUT 2

```
#n4
sentence11 = """
はじめまして、〇〇と申します。インドの大学で経営学を勉強しています。日本語も勉強してい
て、N4レベルです。将来は日本で働きたいと思っています。趣味は読書と映画を見ることです。ど
うぞよろしくお願いします。
"""
```

OUTPUT

Romaji: Hajimemashite,líng líng Tomōshimasu. Indo no daigaku de keiei-gaku o benkyō shite imasu. Nihongo mo benkyō shite ite, N 4 reberudesu. Shōrai wa Nihon de hatarakitai to omotte imasu. Shumi wa dokusho to eiga o miru kotodesu. Dōzoyoroshikuonegaishimasu.

```
{
  "contextual_validity_of_keigo": "Yes",
  "contextual_analysis_of_keigo": "You used the appropriate teineigo form
('~masu') in this introductory context, demonstrating politeness. This is
a common way to introduce yourself and show respect when meeting someone
for the first time.",
  "keigo_count": {
    "teineigo": 1,
    "sonkeigo": 0,
    "kenjougo": 0
  },
  "keigo_analysis": "Your speech demonstrates the use of teineigo, showing
politeness. No sonkeigo (respectful language) detected. No kenjougo
(humble language) detected.  Your keigo usage is basic, focusing primarily
on politeness. Try incorporating respectful and humble forms for more
balanced communication.",
  "proficiency_level": "Basic",
  "improvement_tips": [
    "You could practice using sonkeigo to show respect to those in a
higher position than you.",
    "You could try using kenjougo when talking about yourself or your
actions to show humility."
  ]
}
```

SAMPLE INPUT 3

```
#n2
sentence111 = "はじめまして、〇〇と申します。インドの〇〇大学で経営学を専攻し、今年の
6月に卒業いたしました。在学中は、リーダーシップやコミュニケーション能力を伸ばすために、
学生会やボランティア活動に積極的に参加してまいりました。また、日本語の学習にも力を入れて
おり、現在はJLPT N2を取得しています。日本企業で働くことを目標に、日本とインドのビジネス
関係を深めるための架け橋となりたいと考えております。本日はどうぞよろしくお願いいたしま
す。"
```

OUTPUT

Romaji: Hajimemashite,líng líng Tomōshimasu. Indo no rei rei daigaku de keiei-gaku o senkō shi, kotoshi no 6 tsuki ni sotsugyō itashimashita. Zaigaku-chū wa, rīdāshippu ya komyunikēshon nōryoku o nobasu tame ni, gakusei-kai ya borantia katsudō ni sekkyokutekini sanka shite mairimashita. Mata, nihongo no gakushū ni mo chikara o irete ori, genzai wa JLPT N 2 o shutoku shite imasu. Nihon kigyō de hataraku koto o mokuhyō ni, Nihon to Indo no bijinesu kankei o fukameru tame no kakehashi to naritai to kangaete orimasu. Honjitsu wa dōzo yoroshikuonegaītashimasu.

```
{
  "contextual_validity_of_keigo": "Yes",
  "contextual_analysis_of_keigo": "You used teineigo appropriately to
express politeness, which is suitable for general conversation and when
addressing someone you are not very familiar with. The use of kenjougo
shows humility, which is also appropriate in many situations where you
want to express deference or respect.",
  "keigo_count": {
    "teineigo": 2,
    "sonkeigo": 0,
    "kenjougo": 1
  },
  "keigo_analysis": "Your speech demonstrates the use of teineigo, showing
politeness. No sonkeigo (respectful language) detected. You used kenjougo,
which reflects humility in your speech. This is great for situations where
you need to humble yourself. You are at an intermediate level, showing
both polite and humble expressions. Adding respectful language would help
you communicate with authority figures more effectively.",
  "proficiency_level": "Intermediate",
  "improvement_tips": [
    "You could use more sonkeigo when addressing superiors or people in
positions of authority.",
    "Consider using a combination of teineigo and sonkeigo to express a
deeper level of respect when appropriate."
```

```
    ]}
```

SAMPLE INPUT 4

```
#what makes you ideal for working in our company in japan?


sentence = """
私は日本で御社に貢献できる理想的な候補者だと考えています。まず、私は大学で経営学を専攻し
ており、ビジネスに関する知識を深く学んできました。さらに、日本語能力試験N1を取得してお
り、日本語でのコミュニケーションに自信があります。
御社のビジョンに共感し、企業文化にも深く理解を持っています。今まで、チームでのプロジェク
トマネジメントを経験しており、リーダーシップスキルを培ってまいりました。社長や上司の指示
をしっかりと尊重し、指導をいただいた際には、それをしっかりと反映してまいります。
また、私は多国籍チームで働いた経験があり、異文化理解に長けております。特に、日本のビジネ
スマナーを大切にし、敬語の使い方にも気を付けてまいります。お客様に対しては、丁寧語を使用
し、適切な対応を心掛けております。
さらに、御社の成長に貢献できるよう、これまで学んできた知識と経験を活かし、自ら積極的に動
いて参ります。特に、お客様との商談の際には、謙譲語を使用し、相手を立てることを忘れませ
ん。常に謙虚な姿勢で仕事に取り組み、御社の目標達成に向けて努力いたします。
これらの理由から、私は御社にふさわしい理想的な候補者だと確信しております。ぜひとも、私を
御社で働かせていただける機会をいただければ幸いです。どうぞよろしくお願い申し上げます。
"""
```

OUTPUT

Romaji: Watashi wa Nihon de onsha ni kōken dekiru risō-tekina kōho-shada to kangaete imasu. Mazu, watashi wa daigaku de keiei-gaku o senkō shite ori, bijinesu ni kansuru chishiki o fukaku manande kimashita. Sarani, Nihon gonōryokushiken N 1 o shutoku shite ori, nihongode no komyunikēshon ni jishin ga arimasu. Onsha no bijon ni kyōkan shi, kigyō bunka ni mo fukaku rikai o motte imasu. Ima made, chīmu de no purojekuto manejimento o keiken shite ori, rīdāshippusukiru o tsuchikatte mairimashita. Shachō ya jōshi no shiji o shikkari to sonchō shi, shidō o itadaita sai ni wa, sore o shikkari to han'ei shite mairimasu. Mata, watashi wa ta kokuseki chīmu de hataraita keiken ga ari, ibunkarikai ni takete orimasu. Tokuni, Nihon no bijinesumanā o taisetsu ni shi, keigo no tsukaikata ni mo ki o tsukete mairimasu. Okyakusama ni taishite wa, teinei-go o shiyō shi, tekisetsuna taiō o kokorogakete orimasu. Sarani, onsha no seichō ni kōken dekiru yō, kore made manande kita chishiki to keiken o ikashi, mizukara sekkyokutekini ugoite mairimasu. Tokuni, okyakusama to no shōdan no sai ni wa, kensetsu o shiyō shi, aite o tateru koto o wasuremasen. Tsuneni kenkyona shisei de shigoto ni torikumi, onsha no mokuhyō tassei ni mukete doryoku itashimasu. Korera no riyū kara, watashi wa onsha ni fusawashī risō-tekina kōho-shada to kakushin shite orimasu. Zehitomo, watashi o

onsha de hataraka sete itadakeru kikai o itadakereba saiwaidesu. Dōzo yoroshiku onegai mōshiagemasu.

{
  "contextual_validity_of_keigo": "Yes",
  "contextual_analysis_of_keigo": "You have effectively used a mix of teineigo, sonkeigo, and kenjougo in your speech, demonstrating a strong understanding of the appropriate register for a job application. Your use of teineigo conveys politeness, sonkeigo shows respect towards the company and its representatives, and kenjougo expresses humility, which is essential for a job applicant. ",
  "keigo_count": {
    "teineigo": 16,
    "sonkeigo": 1,
    "kenjougo": 4
  },
  "keigo_analysis": "Your speech demonstrates the use of teineigo, showing politeness. You used sonkeigo, indicating respect towards the listener. This is important when addressing people of higher status. You used kenjougo, which reflects humility in your speech. This is great for situations where you need to humble yourself. Your keigo usage is advanced, showing a well-rounded command of polite, respectful, and humble expressions. Excellent job!",
  "proficiency_level": "Advanced",
  "improvement_tips": [
    "You could experiment with different levels of formality within the same conversation to create a more nuanced and engaging communication style.",
    "Practice using more specific and targeted expressions within each keigo category to refine your communication and showcase your advanced language skills."
  ]
}

SAMPLE INPUT 5

sentence1111=**"「私は今、図書館に行く。」"**

OUTPUT

Romaji: `Watashi wa ima, toshokan ni iku.'

```
{
  "contextual_validity_of_keigo": "No",
  "contextual_analysis_of_keigo": "You have not used any keigo forms in
this sentence. While appropriate for casual conversation, it might be
necessary to use keigo depending on the context and who you are speaking
to. For example, if you are speaking to a teacher or someone of higher
status, using teineigo or sonkeigo would be more appropriate.",
  "keigo_count": {
    "teineigo": 0,
    "sonkeigo": 0,
    "kenjougo": 0
  },
  "keigo_analysis": "No teineigo (polite language) detected. No sonkeigo
(respectful language) detected. No kenjougo (humble language) detected.
No keigo detected. Incorporating polite, respectful, and humble forms will
improve the overall formality and appropriateness of your speech.",
  "proficiency_level": "No keigo detected",
  "improvement_tips": [
    "You could consider using teineigo (polite language) in more formal
situations or when speaking to someone you respect.",
    "You could also learn to use sonkeigo (respectful language) when
talking about someone else, especially in a higher position."
  ]
}
```

EXPECTED OUTPUTS:

More changes shall be added later on the basis of benchmarks.

QUESTION-WISE KEIGO ANALYSIS

✓ HUMBLENESS

✗ POLITENESS

✗ RESPECTFULNESS

✓ CONTEXTUAL CORRECTNESS

IMPROVEMENT AREAS

- YOU SHOULD USE MORE HUMBLE FORMS WHEN SPEAKING ABOUT YOURSELF IN FORMAL SETTINGS, ESPECIALLY WHEN ADDRESSING SUPERIORS.
- YOU COULD PRACTICE USING SONKEIGO AND KENJOUGO IN DIFFERENT SITUATIONS TO ENHANCE YOUR KEIGO USAGE.

OVERALL KEIGO ANALYSIS

PROFICIENCY LEVEL

BASIC

CONTEXTUAL ANALYSIS

YOUR OVERALL ANSWERS HAVE CORRECT CONTEXTUAL AWARENESS OF KEIGOS. YOU USED THE APPROPRIATE TEINEIGO FORMS FOR THE GIVEN CONTEXT, WHICH IS A SELF-INTRODUCTION IN A FORMAL SETTING.

OVERALL IMPROVEMENT AREAS

- CONSIDER USING RESPECTFUL LANGUAGE WHEN ADDRESSING SUPERIORS, LIKE 'IRASSHAIMASU' FOR 'TO GO' OR 'NASAIMASU' FOR 'TO DO'.
- TRY USING HUMBLE FORMS LIKE 'MOUSHIMASU' FOR 'TO SAY' OR 'ITADAKIMASU' FOR 'TO RECEIVE' WHEN SPEAKING ABOUT YOUR ACTIONS IN A FORMAL SETTING.
- TRY INCORPORATING RESPECTFUL AND HUMBLE FORMS FOR MORE BALANCED COMMUNICATION.