the sample size was 104 . definitely more samples caused more accuracy. and in this short time , it wasn't possible.

Please do not copy this code directly. Use it as a reference for learning and write your own implementation

Section1

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import LabelEncoder, StandardScaler
from sklearn.preprocessing import MinMaxScaler
```

```
# Load the data
data = pd.read_csv('/content/Online questionnaire _ (Responses) - Form Responses 5.csv')
print(data.columns)
```

```
Index(['Timestamp', '1.Age', '2.Gender', '3.Country/Region',
       '4.Field of Study', '5.Level of Education',
       '6.Are you currently employed?',
       '7.If employed, what is your current role?',
       '8.If not employed, how many job vacancies have you applied for in the past 6 months?',
       '9.How familiar are you with AI tools (e.g., ChatGPT, LinkedIn AI resume builders, career prediction tools)?',
       '10. How frequently do you use AI tools for job-related purposes?',
       '11.For which tasks have you used AI tools?(Check all that apply)',
       '12.Which AI tools do you use most frequently ?',
       '13.Do you think  AI tools make you ..... in job related task. (select all appropriate options if applied)',
       '14.On a scale of 1 to 5 , (1 = Strongly Disagree,5 = Strongly Agree)\n\nrate the following statements about your emotions when using AI tools:\n\n   14.1) I feel motivated to apply for jobs after using
       '14.2)  Using AI tools makes me feel less stressed about job hunting.  \n',
       '14.3) I feel unsure about trusting AI tools’ recommendations.\n',
       '14.4) AI tools give me a sense of control over my job search process and an application.  ',
       '14.5) Using AI tools makes me feel alienated from traditional job-seeking methods.  ',
       '15.How would you compare the emotional support provided by AI tools to human advisors (e.g., career counselors)',
       '16. Can you describe a situation where an AI tool alleviated or intensified your stress during the job search process , during job interview ?',
       '17.1) AI tools make me feel more competitive compared to other candidates.  ',
       '17.2) AI tools reinforce my belief in my skills and abilities.\n',
       '17.2) AI tools sometimes make me doubt my ability to perform tasks without assistance.\n',
       '17.3) Relying on AI tools may reduce my individuality in job applications\n',
       '18.How confident are you in your ability to succeed in the job market with the help of AI tools?',
       '19.  Have you ever felt excluded or disadvantaged in the job market due to limited access to advanced AI tools? Why or why not?',
       '20.How have AI tools impacted your job performance?',
       '21.If positively, then describe how AI tools contributed to your job satisfaction and performance .',
       '22.If negatively, describe any challenges or frustrations you’ve faced with AI tools.',
       '23.1) I feel confidential when I secure a job with AI tools',
       '23.2) AI tools align with my career aspirations.\n',
       '23.3) I feel a lack of personal connection in AI-assisted job processes.\n',
       '23.4) Using AI tools impact my emotional well-being during job applications\n',
       '24.How do you perceive AI tools’ role in shaping the future job market?',
       '25.Do you think AI tools have reduced or increased competition in the job market? Why?',
       '26.How do you think employers view candidates who rely on AI tools for job search and applications?',
       '27.What features would you like to see in AI tools to improve the job search experience?',
       '28.Should universities or career services provide training on using AI tools for job searches?\n Why or why not?',
       '29.  What should the designers of AI-based tools apply to motivate and help the job seekers on their search and resume?',
       ' Your support would be greatly appreciated and will significantly enhance the quality of the research and this part is completely optional. 🌟\nThank you once again for your time! 🌻 \n'],
      dtype='object')
```

```
print(data.shape)
```

```
(102, 41)
```

```
# Rename columns to match the desired format
data.rename(columns={
    '1.Age': 'Age',
    '2.Gender': 'Gender',
    '3.Country/Region': 'Country/Region',
    '4.Field of Study': 'Field of Study',
    '5.Level of Education': 'Level of Education',
    '6.Are you currently employed?': 'Are you currently employed?',
    '7.If employed, what is your current role?': 'If employed, what is your current role?',
    '8.If not employed, how many job vacancies have you applied for in the past 6 months?': 'Job applications in the past 6 months'
}, inplace=True)

# Check if the columns were renamed successfully
print(data.columns)


# Select only relevant columns for Section 1
data = data[['Age', 'Gender', 'Country/Region', 'Field of Study', 'Level of Education', 'Are you currently employed?', 'If employed, what is your current role?', 'Job applications in the past 6 months']]

# Handle missing values
data['Age'] = pd.to_numeric(data['Age'], errors='coerce')  # Convert age to numeric, force errors to NaN
data['Age'].fillna(data['Age'].median(), inplace=True)  # Fill missing age with median

# Fill missing text fields with 'Unknown'
text_columns = ['Gender', 'Country/Region', 'Field of Study', 'Level of Education', 'Are you currently employed?', 'If employed, what is your current role?', 'Job applications in the past 6 months']
for col in text_columns:
    data[col] = data[col].fillna('Unknown')
```

```
Index(['Timestamp', 'Age', 'Gender', 'Country/Region', 'Field of Study',
       'Level of Education', 'Are you currently employed?',
       'If employed, what is your current role?',
       'Job applications in the past 6 months',
       '9.How familiar are you with AI tools (e.g., ChatGPT, LinkedIn AI resume builders, career prediction tools)?',
       '10. How frequently do you use AI tools for job-related purposes?',
       '11.For which tasks have you used AI tools?(Check all that apply)',
       '12.Which AI tools do you use most frequently ?',
       '13.Do you think  AI tools make you ..... in job related task. (select all appropriate options if applied)',
       '14.On a scale of 1 to 5 , (1 = Strongly Disagree,5 = Strongly Agree)\n\nrate the following statements about your emotions when using AI tools:\n\n   14.1) I feel motivated to apply for jobs after using
       '14.2)  Using AI tools makes me feel less stressed about job hunting.  \n',
       '14.3) I feel unsure about trusting AI tools’ recommendations.\n',
       '14.4) AI tools give me a sense of control over my job search process and an application.  ',
       '14.5) Using AI tools makes me feel alienated from traditional job-seeking methods.  ',
       '15.How would you compare the emotional support provided by AI tools to human advisors (e.g., career counselors)',
       '16. Can you describe a situation where an AI tool alleviated or intensified your stress during the job search process , during job interview ?',
       '17.1) AI tools make me feel more competitive compared to other candidates.  ',
       '17.2) AI tools reinforce my belief in my skills and abilities.\n',
       '17.2) AI tools sometimes make me doubt my ability to perform tasks without assistance.\n',
       '17.3) Relying on AI tools may reduce my individuality in job applications\n',
       '18.How confident are you in your ability to succeed in the job market with the help of AI tools?',
       '19.  Have you ever felt excluded or disadvantaged in the job market due to limited access to advanced AI tools? Why or why not?',
       '20.How have AI tools impacted your job performance?',
       '21.If positively, then describe how AI tools contributed to your job satisfaction and performance .',
```

```
        '22.If negatively, describe any challenges or frustrations you've faced with AI tools.',
        '23.1) I feel confidential when I secure a job with AI tools',
        '23.2) AI tools align with my career aspirations.\n',
        '23.3) I feel a lack of personal connection in AI-assisted job processes.\n',
        '23.4) Using AI tools impact my emotional well-being during job applications\n',
        '24.How do you perceive AI tools' role in shaping the future job market?',
        '25.Do you think AI tools have reduced or increased competition in the job market? Why?',
        '26.How do you think employers view candidates who rely on AI tools for job search and applications?',
        '27.What features would you like to see in AI tools to improve the job search experience?',
        '28.Should universities or career services provide training on using AI tools for job searches?\n Why or why not?',
        '29.  What should the designers of AI-based tools apply to motivate and help the job seekers on their search and resume?',
        ' Your support would be greatly appreciated and will significantly enhance the quality of the research and this part is completely optional. 🌟\nThank you once again for your time! 🌻 \n'],
      dtype='object')
<ipython-input-5-59d198960a36>:22: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original obje


  data['Age'].fillna(data['Age'].median(), inplace=True)  # Fill missing age with median
```

```
# Clean Country/Region (make consistent capitalization)
data['Country/Region'] = data['Country/Region'].str.title()

# Clean Field of Study
field_mapping = {
    'it': 'Information Technology',
    'iit': 'Information Technology',
    'information technology': 'Information Technology',
    'information of technology': 'Information Technology',
    'ict': 'Information Technology',
    'ict engineering': 'Information Technology',
    'ict software engineering': 'Information Technology',
    'ict software': 'Information Technology',
    'information and communication technology': 'Information Technology',
    'information and communications technology (smart iot systems)': 'Information Technology',
    'computer science': 'Computer Science',
    'computer': 'Computer Science',
    'software engineering': 'Software Engineering',
    'software engineer': 'Software Engineering',
    'sofware engineering': 'Software Engineering',
    'software development': 'Software Engineering',
    'ai': 'Artificial Intelligence',
    'digital transformation': 'Digital Transformation',
    'business information technology': 'Business Analytics',
    'business analytics': 'Business Analytics',
    'mba': 'Business Administration',
    'bba(hrm)': 'Business Administration',
    'finance': 'Finance',
    'economic': 'Economics',
    'nursing': 'Nursing',
    'biology': 'Biology',
    'architecture': 'Architecture',
    'industrial management': 'Industrial Management',
    'environmental engineering': 'Environmental Engineering',
    'electrical engineering': 'Electrical Engineering',
    'electronic engineering': 'Electronic Engineering',
    'media management': 'Media Management',
    'management': 'Management',
```

```
    'cancer medicine': 'Cancer Medicine',
    'animal science': 'Animal Science',
    'peace, mediation and conflict research': 'Peace and Conflict Research',
    'engineering': 'Engineering',
    'medical technology': 'Medical Technology',
    'registred nurse (master)': 'Nursing',
    'masterâ€™s': 'Masters Degree',
    "master's degree programme in information technology": 'Masters Degree',
    "master of information technology": 'Masters Degree',
    "master of engineering, medical technology": 'Medical Technology'
}

# Normalize Field of Study
def normalize_field(field):
    field = str(field).lower().strip()
    return field_mapping.get(field, field.title())

data['Field of Study'] = data['Field of Study'].apply(normalize_field)
```

```
# Correct If not employed, how many job vacancies
data['Job applications in the past 6 months'] = pd.to_numeric(data['Job applications in the past 6 months'], errors='coerce')
data['Job applications in the past 6 months'].fillna(0, inplace=True)

# Convert 'Age' to numeric, force errors to NaN
data['Age'] = pd.to_numeric(data['Age'], errors='coerce')

# Fill missing 'Age' values with the median
data['Age'].fillna(data['Age'].median(), inplace=True)

# Encode categorical columns
categorical_cols = ['Gender', 'Country/Region', 'Field of Study', 'Level of Education', 'Are you currently employed?']
label_encoders = {}
for col in categorical_cols:
    le = LabelEncoder()
    data[col] = le.fit_transform(data[col])
    label_encoders[col] = le
```

```
<ipython-input-7-3a358f30ac2f>:3: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original obje


  data['Job applications in the past 6 months'].fillna(0, inplace=True)
<ipython-input-7-3a358f30ac2f>:9: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original obje


  data['Age'].fillna(data['Age'].median(), inplace=True)
```

```
# Check the unique values in the 'Age' column to understand the data distribution
print(data['Age'].unique())

# Filter any negative or unrealistic values for Age, just in case
data['Age'] = data['Age'].apply(lambda x: x if x >= 0 and x <= 100 else None)
```

```
# Fill missing values with the median again after fixing
data['Age'] = data['Age'].fillna(data['Age'].median())  # Replace missing values with the median of the 'Age' column

# Convert 'Age' to numeric (if necessary)
data['Age'] = pd.to_numeric(data['Age'], errors='coerce')  # Convert 'Age' to numeric, forcing errors to NaN

# Check again after filling missing values
print(data['Age'].describe())
```

```
[25. 54. 38. 29. 22. 37. 36. 35. 30. 34. 33. 32. 31. 27. 39. 42. 41. 28.
 23. 20. 40. 26. 21. 45. 44. 24. 43. 47.]
count    102.000000
mean      31.960784
std        6.744153
min       20.000000
25%       27.000000
50%       32.000000
75%       36.000000
max       54.000000
Name: Age, dtype: float64
```

```
# Print Country/Region mapping
print("Country/Region mapping:")
for idx, value in enumerate(label_encoders['Country/Region'].classes_):
    print(f"{idx} -> {value}")

print("\nField of Study mapping:")
# Print Field of Study mapping
for idx, value in enumerate(label_encoders['Field of Study'].classes_):
    print(f"{idx} -> {value}")
```

```
7 -> Finland
8 -> Finland
9 -> Finland Uusimaa
10 -> Finland, Uusimaa
11 -> Finland/ Uusima
12 -> Finland/Espoo
13 -> Finland/Turku
14 -> Finland/Uusimaa
15 -> Germany
16 -> Helsinki
17 -> Helsinki, Finland
18 -> India
19 -> Iran
20 -> Iran
21 -> Iran/Tehran
22 -> Malaysia
23 -> Nationality:Chinese; Live In Finland Now
24 -> Nigeria
25 -> Portugal
26 -> Switzerland
27 -> United Kingdom
```

```
4 -> Business Administration
5 -> Business Analytics
6 -> Cancer Medicine
7 -> Computer Engineering
8 -> Computer Science
9 -> Digitale Transformation
10 -> Economics
11 -> Electrical Engineering
12 -> Electronic Engineering
13 -> Engineering
14 -> Environmental Engineering
15 -> Finance
16 -> Industrial Management
17 -> Information And Communication Technologies
18 -> Information And Communications Techonology (Smart Iot Systems)
19 -> Information Technology
20 -> Information Technology, Masters
21 -> It Engineering
22 -> Managment
23 -> Masters Degree
24 -> Master’S
25 -> Media Management
26 -> Medical Technology
27 -> Nursing
28 -> Peace and Conflict Research
29 -> Software Engineer, Iot Embedded Devices
30 -> Software Engineering
31 -> Software Engineering/ Ai
32 -> Swe
```

Section2

```
#  Check the first few rows of the data
print(data.head())

#  Check data types of each column
print(data.dtypes)

# Check the number of rows and columns
print(data.shape)

# Remove any extra spaces from column names
data.columns = data.columns.str.strip()

# Specific Cleaning for Questions 9-12

# List of actual columns related to questions 9-12
text_columns = [
    'Familiarity with AI tools',
    'Frequency of AI tool usage for job purposes',
    'Tasks used AI tools for',
    'Most frequently used AI tools'
]

for col in text_columns:
    if col in data.columns:
        # If the column was text and empty, fill it with "Unknown"
        if data[col].dtype == 'object':
            data[col] = data[col].fillna('Unknown')
        else:
```

```
          # If it was a number, fill in with a numerical average
            data[col] = data[col].fillna(data[col].mean())

#  Drop completely empty rows (still needed)
#data = data.dropna(how='all')

# Reset index after cleaning
data = data.reset_index(drop=True)

#  Check the data again after cleaning
print(data.head())
print(data.shape)
```

```
            Timestamp 1.Age 2.Gender 3.Country/Region         4.Field of Study  \
0   1/14/2025 9:37    25   Female          finland  Information technology
1  1/19/2025 15:14    54     Male          Finland                      IT
2  1/19/2025 15:26    38     Male         Finland                       IT
3  1/19/2025 15:38    29   Female         Finland   Information technology
4  1/19/2025 16:04    22     Male        Malaysia         Computer Science

  5.Level of Education 6.Are you currently employed?  \
0         Postgraduate                           No
1         Postgraduate               Yes (Full-time)
2             Graduate               Yes (Full-time)
3        Undergraduate               Yes (Full-time)
4        Undergraduate                           No

  7.If employed, what is your current role?  \
0                                       NaN
1                             Security Lead
2                         Testing specialist
3                           Senior developer
4                                       NaN

  8.If not employed, how many job vacancies have you applied for in the past 6 months?  \
0                                                 40
1                                                NaN
2                                                NaN
3                                                NaN
4                                                  3

  9.How familiar are you with AI tools (e.g., ChatGPT, LinkedIn AI resume builders, career prediction tools)?  \
0                                    Very familiar
1                                 Somewhat familiar
2                                           Neutral
3                                 Somewhat familiar
4                                    Very familiar

   ... 23.2) AI tools align with my career aspirations.\n  \
0  ...                                             NaN
1  ...                                             4.0
2  ...                                             4.0
3  ...                                             4.0
4  ...                                             5.0

  23.3) I feel a lack of personal connection in AI-assisted job processes.\n  \
0                                                NaN
1                                                3.0
2                                                4.0
3                                                3.0
4                                                3.0
```

```
  23.4) Using AI tools impact my emotional well-being during job applications\n  \
0                                                NaN
1                                                2.0
2                                                3.0
3                                                3.0
4                                                3.0
```

```
# Rename columns to match the desired format for Section 2
data.rename(columns={
    '9.How familiar are you with AI tools (e.g., ChatGPT, LinkedIn AI resume builders, career prediction tools)?': 'Familiarity with AI tools',
    '10. How frequently do you use AI tools for job-related purposes?': 'Frequency of AI tool usage for job purposes',
    '11.For which tasks have you used AI tools?(Check all that apply)': 'Tasks used AI tools for',
    '12.Which AI tools do you use most frequently ?': 'Most frequently used AI tools',
    '13.Do you think  AI tools make you ..... in job related task. (select all appropriate options if applied)': 'Impact of AI tools on job tasks',
    '14.On a scale of 1 to 5 , (1 = Strongly Disagree,5 = Strongly Agree)\n\nrate the following statements about your emotions when using AI tools:\n\n   14.1) I feel motivated to apply for jobs after using AI tool
    '14.2)  Using AI tools makes me feel less stressed about job hunting.': 'Less stress using AI tools for job hunting',
    '14.3) I feel unsure about trusting AI tools' recommendations.': 'Uncertainty about trusting AI recommendations',
    '14.4) AI tools give me a sense of control over my job search process and an application.': 'Sense of control in job search using AI tools',
    '14.5) Using AI tools makes me feel alienated from traditional job-seeking methods.': 'Alienation from traditional job-seeking methods using AI tools',
    '15.How would you compare the emotional support provided by AI tools to human advisors (e.g., career counselors)': 'Emotional support comparison: AI tools vs human advisors',
    '16. Can you describe a situation where an AI tool alleviated or intensified your stress during the job search process , during job interview ?': 'AI tool impact on stress during job search or interview',
    '17.On a scale of 1 to 5 (1 = Strongly Disagree, 5 = Strongly Agree),\n rate how AI tools influence your self-perception in the job market:\n\n17.1) AI tools make me feel more competitive compared to other cand
    '17.2) AI tools reinforce my belief in my skills and abilities.': 'AI tools reinforcing belief in skills',
    '17.2) AI tools sometimes make me doubt my ability to perform tasks without assistance.': 'Doubts about ability without AI assistance',
    '17.3) Relying on AI tools may reduce my individuality in job applications': 'AI tools reducing individuality in job applications',
    '18.How confident are you in your ability to succeed in the job market with the help of AI tools?': 'Confidence in job market success with AI tools',
    '19.  Have you ever felt excluded or disadvantaged in the job market due to limited access to advanced AI tools? Why or why not?': 'Feeling excluded due to limited AI access',
    '20.How have AI tools impacted your job performance?': 'Impact of AI tools on job performance',
    '21.If positively, then describe how AI tools contributed to your job satisfaction and performance .': 'Positive impact of AI tools on satisfaction and performance',
    '22.If negatively, describe any challenges or frustrations you've faced with AI tools.': 'Negative impact of AI tools: challenges and frustrations',
    '23.On a scale of 1 to 5 (1 = Strongly Disagree, 5 = Strongly Agree), \nrate the following statements about job satisfaction and emotions: \n\n23.1) I feel confidential when I secure a job with AI tools help.':
    '23.2) AI tools align with my career aspirations.': 'AI tools alignment with career aspirations',
    '23.3) I feel a lack of personal connection in AI-assisted job processes.': 'Lack of personal connection in AI job processes',
    '23.4) Using AI tools impact my emotional well-being during job applications': 'Emotional impact of AI tools in job applications',
    '24.How do you perceive AI tools' role in shaping the future job market?': 'AI tools role in shaping future job market',
    '25.Do you think AI tools have reduced or increased competition in the job market? Why?': 'Impact of AI tools on competition in the job market',
    '26.How do you think employers view candidates who rely on AI tools for job search and applications?': 'Employer views on AI-reliant candidates',
    '27.What features would you like to see in AI tools to improve the job search experience?': 'Desired features in AI tools for job search',
    '28.Should universities or career services provide training on using AI tools for job searches?\n Why or why not?': 'Training on AI tools for job search',
    '29.  What should the designers of AI-based tools apply to motivate and help the job seekers on their search and resume?': 'Design suggestions for AI tools to assist job seekers'
}, inplace=True)
```

```
print(data.head())
print(data.shape)
```

```
           Timestamp 1.Age 2.Gender 3.Country/Region         4.Field of Study  \
0   1/14/2025 9:37    25   Female          finland  Information technology
1  1/19/2025 15:14    54     Male          Finland                      IT
2  1/19/2025 15:26    38     Male         Finland                       IT
3  1/19/2025 15:38    29   Female         Finland   Information technology
4  1/19/2025 16:04    22     Male        Malaysia         Computer Science

  5.Level of Education 6.Are you currently employed?  \
0         Postgraduate                           No
1         Postgraduate              Yes (Full-time)
2             Graduate              Yes (Full-time)
3        Undergraduate              Yes (Full-time)
```

```
4       Undergraduate                            No

  7.If employed, what is your current role?  \
0                                        NaN
1                             Security Lead
2                          Testing specialist
3                            Senior developer
4                                        NaN

  8.If not employed, how many job vacancies have you applied for in the past 6 months?  \
0                                                 40
1                                                NaN
2                                                NaN
3                                                NaN
4                                                  3

  Familiarity with AI tools  ... AI tools alignment with career aspirations  \
0            Very familiar  ...                                        NaN
1        Somewhat familiar  ...                                        4.0
2                  Neutral  ...                                        4.0
3        Somewhat familiar  ...                                        4.0
4            Very familiar  ...                                        5.0

  Lack of personal connection in AI job processes  \
0                                             NaN
1                                             3.0
2                                             4.0
3                                             3.0
4                                             3.0

  Emotional impact of AI tools in job applications  \
0                                              NaN
1                                              2.0
2                                              3.0
3                                              3.0
4                                              3.0

         AI tools role in shaping future job market  \
0                                               NaN
1  It is just a new tool. The hype is huge around...
2                   People will be replaced with AI
3  Experience in efficient use of AI tools will d...
4                                               NaN

   Impact of AI tools on competition in the job market  \
0                                               NaN
```

```
# Print data after renaming columns
print(data.head())
print(data.shape)

# Check if the columns still exist and the changes were made correctly
print(data.columns)
```

```
0        Postgraduate                            No
1        Postgraduate              Yes (Full-time)
```

```
5                          Senior developer
4                                       NaN

  8.If not employed, how many job vacancies have you applied for in the past 6 months?  \
0                                                 40
1                                                NaN
2                                                NaN
3                                                NaN
4                                                  3

  Familiarity with AI tools  ... AI tools alignment with career aspirations  \
0            Very familiar  ...                                        NaN
1        Somewhat familiar  ...                                        4.0
2                  Neutral  ...                                        4.0
3        Somewhat familiar  ...                                        4.0
4            Very familiar  ...                                        5.0

  Lack of personal connection in AI job processes  \
0                                             NaN
1                                             3.0
2                                             4.0
3                                             3.0
4                                             3.0

  Emotional impact of AI tools in job applications  \
0                                              NaN
1                                              2.0
2                                              3.0
3                                              3.0
4                                              3.0

         AI tools role in shaping future job market  \
0                                              NaN
1  It is just a new tool. The hype is huge around...
2                   People will be replaced with AI
3  Experience in efficient use of AI tools will d...
4                                              NaN

   Impact of AI tools on competition in the job market  \
0                                              NaN
1  It will change the competition, but I don't th...
2                                         Increased
3  AI is a tool and and as every tool you need to...
4                                              NaN

            Employer views on AI-reliant candidates  \
0                                              NaN
1  I don't have any opinion. I guess it will vary...
```

```
# Remove leading/trailing spaces from column names
data.columns = data.columns.str.strip()

#  Rename columns based on the actual names
data.rename(columns={
    '9.How familiar are you with AI tools (e.g., ChatGPT, LinkedIn AI resume builders, career prediction tools)?': 'Familiarity with AI tools',
    '10. How frequently do you use AI tools for job-related purposes?': 'Frequency of AI tool usage for job purposes',
    '11.For which tasks have you used AI tools?(Check all that apply)': 'Tasks used AI tools for',
    '12.Which AI tools do you use most frequently ?': 'Most frequently used AI tools'
}, inplace=True)


for column in ['Familiarity with AI tools', 'Frequency of AI tool usage for job purposes', 'Tasks used AI tools for', 'Most frequently used AI tools']:
    print(data[column].head(), '\n')
```

```
0       Very familiar
1    Somewhat familiar
2              Neutral
3    Somewhat familiar
4       Very familiar
Name: Familiarity with AI tools, dtype: object

0     Weekly
1     Weekly
2     Rarely
3    Monthly
4      Daily
Name: Frequency of AI tool usage for job purposes, dtype: object

0    Resume and cover letter creation, Job search a...
1                                     Grammar checking
2                                       I haven't used
3    Assisting in information searching and process...
4    Resume and cover letter creation, Job search a...
Name: Tasks used AI tools for, dtype: object

0                          chatgpt
1    Grammarly, Gemini, ChatGPT,...
2                           Copilot
3                           ChatGPT
4                          ChatGPT
Name: Most frequently used AI tools, dtype: object
```

```
# Data Cleaning (Converting to lowercase and stripping extra spaces)
data['Familiarity with AI tools'] = data['Familiarity with AI tools'].str.strip().str.lower()
data['Frequency of AI tool usage for job purposes'] = data['Frequency of AI tool usage for job purposes'].str.strip().str.lower()
data['Tasks used AI tools for'] = data['Tasks used AI tools for'].str.strip().str.lower()
data['Most frequently used AI tools'] = data['Most frequently used AI tools'].str.strip().str.lower()

# Mapping values to numerical equivalents
familiarity_mapping = {
    'very familiar': 5,
    'somewhat familiar': 4,
    'neutral': 3,
    'slightly familiar': 2,
    'not familiar at all': 1
}
data['Familiarity with AI tools'] = data['Familiarity with AI tools'].map(familiarity_mapping).fillna(0).astype(int)

frequency_mapping = {
    'daily': 5,
    'weekly': 4,
    'monthly': 3,
    'rarely': 2,
    'never': 1
}
data['Frequency of AI tool usage for job purposes'] = data['Frequency of AI tool usage for job purposes'].map(frequency_mapping).fillna(0).astype(int)

tasks_mapping = {
    'resume and cover letter creation': 1,
    'job search and application tracking': 2,
    'interview preparation': 3,
    'skill assessment and career advice': 4,
    'networking assistance': 5,
    'other': 6
```

```
}
data['Tasks used AI tools for'] = data['Tasks used AI tools for'].map(tasks_mapping).fillna(0).astype(int)

ai_tools_mapping = {
    'chatgpt': 1,
    'gemini': 2,
    'copilot': 3,
    'other': 4
}
data['Most frequently used AI tools'] = data['Most frequently used AI tools'].map(ai_tools_mapping).fillna(0).astype(int)
```

```
# Rename columns to match the desired format for Section 2
data.rename(columns={
    '9.How familiar are you with AI tools (e.g., ChatGPT, LinkedIn AI resume builders, career prediction tools)?': 'Familiarity with AI tools',
    '10. How frequently do you use AI tools for job-related purposes?': 'Frequency of AI tool usage for job purposes',
    '11.For which tasks have you used AI tools?(Check all that apply)': 'Tasks used AI tools for',
    '12.Which AI tools do you use most frequently ?': 'Most frequently used AI tools',

}, inplace=True)
```

```
# Checking for any remaining NaN values
print(data[['Familiarity with AI tools', 'Frequency of AI tool usage for job purposes']].isna().sum())
print(data[['Familiarity with AI tools', 'Frequency of AI tool usage for job purposes']].head())
```

```
Familiarity with AI tools                    0
Frequency of AI tool usage for job purposes    0
dtype: int64
   Familiarity with AI tools  Frequency of AI tool usage for job purposes
0                          5                                            4
1                          4                                            4
2                          3                                            2
3                          4                                            3
4                          5                                            5
```

```
# Rename columns to match the desired format for Section 2
data.rename(columns={
    '9.How familiar are you with AI tools (e.g., ChatGPT, LinkedIn AI resume builders, career prediction tools)?': 'Familiarity with AI tools',
    '10. How frequently do you use AI tools for job-related purposes?': 'Frequency of AI tool usage for job purposes',
    '11.For which tasks have you used AI tools?(Check all that apply)': 'Tasks used AI tools for',
    '12.Which AI tools do you use most frequently ?': 'Most frequently used AI tools',

}, inplace=True)
```

```
#  Scaling data (if necessary)
scaler = StandardScaler()
data[['Familiarity with AI tools', 'Frequency of AI tool usage for job purposes']] = scaler.fit_transform(
    data[['Familiarity with AI tools', 'Frequency of AI tool usage for job purposes']]
)

# Display the data after preprocessing
print(data.head())
```

```
          Timestamp 1.Age 2.Gender 3.Country/Region         4.Field of Study  \
0   1/14/2025 9:37    25   Female          finland  Information technology
1  1/19/2025 15:14    54     Male          Finland                      IT
2  1/19/2025 15:26    38     Male         Finland                       IT
```

```
3  1/19/2025 15:38    29   Female         Finland   Information technology
4  1/19/2025 16:04    22     Male        Malaysia         Computer Science

  5.Level of Education 6.Are you currently employed?  \
0         Postgraduate                            No
1         Postgraduate              Yes (Full-time)
2             Graduate              Yes (Full-time)
3        Undergraduate              Yes (Full-time)
4        Undergraduate                            No

  7.If employed, what is your current role?  \
0                                       NaN
1                             Security Lead
2                         Testing specialist
3                           Senior developer
4                                       NaN

  8.If not employed, how many job vacancies have you applied for in the past 6 months?  \
0                                                 40
1                                                NaN
2                                                NaN
3                                                NaN
4                                                  3

   Familiarity with AI tools  ...  AI tools alignment with career aspirations  \
0                   0.817053  ...                                        NaN
1                  -0.373510  ...                                        4.0
2                  -1.564072  ...                                        4.0
3                  -0.373510  ...                                        4.0
4                   0.817053  ...                                        5.0

   Lack of personal connection in AI job processes  \
0                                              NaN
1                                              3.0
2                                              4.0
3                                              3.0
4                                              3.0

   Emotional impact of AI tools in job applications  \
0                                               NaN
1                                               2.0
2                                               3.0
3                                               3.0
4                                               3.0

         AI tools role in shaping future job market  \
0                                               NaN
1  It is just a new tool. The hype is huge around...
2                   People will be replaced with AI
3  Experience in efficient use of AI tools will d...
4                                               NaN

   Impact of AI tools on competition in the job market  \
0                                               NaN
```

section 3

```
#  Question 13: One-hot encode multi-select column

## Then clean and one-hot encode
data['Impact_clean'] = data['Impact of AI tools on job tasks'].str.lower().str.strip()
```

```
emotion_options = ['Empowered', 'Confident', 'Anxious', 'Frustrated', 'Overwhelmed', 'Hopeful', 'Other']
for emotion in emotion_options:
    data[emotion] = data['Impact_clean'].str.contains(emotion.lower(), na=False).astype(int)


#  Questions 14.1 to 14.5: Likert scale questions (1 to 5)

likert_cols = [
    'Motivation to apply for jobs using AI tools',
    'Less stress using AI tools for job hunting',
    'Uncertainty about trusting AI recommendations',
    'Sense of control in job search using AI tools',
    'Alienation from traditional job-seeking methods using AI tools'
]

# Convert responses to numeric and fill missing values with column mean
for col in likert_cols:
    data[col] = pd.to_numeric(data[col], errors='coerce')  # Convert to numbers
    data[col].fillna(data[col].mean(), inplace=True)       # Fill missing with mean

# Question 15: Ordinal mapping (categorical to numeric)

support_map = {
    'Much worse': 1,
    'Somewhat worse': 2,
    'About the same': 3,
    'Somewhat better': 4,
    'Much better': 5
    }
```

```
<ipython-input-33-32c97a28f943>:24: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original obje


  data[col].fillna(data[col].mean(), inplace=True)       # Fill missing with mean
```

```
# Map responses and fill missing with mode
data['Emotional support comparison: AI tools vs human advisors'] = data['Emotional support comparison: AI tools vs human advisors'].map(support_map)
data['Emotional support comparison: AI tools vs human advisors'].fillna(data['Emotional support comparison: AI tools vs human advisors'].mode()[0], inplace=True)
```

```
<ipython-input-35-5312202dbaa4>:3: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original obje


  data['Emotional support comparison: AI tools vs human advisors'].fillna(data['Emotional support comparison: AI tools vs human advisors'].mode()[0], inplace=True)
```

section 4

```
data.rename(columns={
    '17.1) AI tools make me feel more competitive compared to other candidates.': 'Feeling more competitive using AI tools',
```

```
    'AI tools reinforcing belief in skills': 'AI tools reinforcing belief in skills',
    'Doubts about ability without AI assistance': 'Doubts about ability without AI assistance',
    'AI tools reducing individuality in job applications': 'AI tools reducing individuality in job applications',
    '18.How confident are you in your ability to succeed in the job market with the help of AI tools?': 'Confidence in job market success with AI tools'
}, inplace=True)

# 1. Define Likert-scale columns from Q17
likert_cols_17 = [
    'Feeling more competitive using AI tools',
    'AI tools reinforcing belief in skills',
    'Doubts about ability without AI assistance',
    'AI tools reducing individuality in job applications'
]

# 2. Convert Likert-scale answers to numeric (1–5)
for col in likert_cols_17:
    data[col] = pd.to_numeric(data[col], errors='coerce')
    data[col].fillna(data[col].mean(), inplace=True)

# 3. Convert Q18 (Confidence in job market success) to numeric
confidence_mapping = {
    'Very doubtful': 1,
    'Slightly doubtful': 2,
    'Neutral': 3,
    'Somewhat confident': 4,
    'Very confident': 5
}
data['Confidence in job market success with AI tools'] = data[
    'Confidence in job market success with AI tools'
].map(confidence_mapping)

# Optional: fill missing with mode (most common answer)
data['Confidence in job market success with AI tools'].fillna(
    data['Confidence in job market success with AI tools'].mode()[0], inplace=True
)
```

```
<ipython-input-38-76dd506c32dd>:20: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original obje


  data[col].fillna(data[col].mean(), inplace=True)
<ipython-input-38-76dd506c32dd>:35: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original obje


  data['Confidence in job market success with AI tools'].fillna(
```

Double-click (or enter) to edit

section 5

```
# Clean up column names: remove leading/trailing spaces and newlines
data.columns = data.columns.str.strip()
data.columns = data.columns.str.replace('\n', '', regex=True)

# Rename columns for easier reference (Section 20 & Section 23)
data.rename(columns={
    '20.How have AI tools impacted your job performance?': 'Impact of AI tools on job performance',
    '23.1) I feel confidential when I secure a job with AI tools': 'Confidence in job security with AI tools',
    '23.2) AI tools align with my career aspirations.': 'AI tools alignment with career aspirations',
    '23.3) I feel a lack of personal connection in AI-assisted job processes.': 'Lack of personal connection in AI job processes',
    '23.4) Using AI tools impact my emotional well-being during job applications': 'Emotional impact of AI tools in job applications'
}, inplace=True)

# Define column lists for Section 5
section5_cols23 = [
    'Confidence in job security with AI tools',
    'AI tools alignment with career aspirations',
    'Lack of personal connection in AI job processes',
    'Emotional impact of AI tools in job applications'
]

section5_cols20 = ['Impact of AI tools on job performance']

# Define the impact_mapping dictionary
impact_mapping = {
    'No impact': 0,
    'Positively': 1,
    'Negatively': -1,
}

# Fill missing values with 0 (representing 'No impact') for the 'Impact of AI tools on job performance' column
data['Impact of AI tools on job performance'] = data['Impact of AI tools on job performance'].fillna('No impact')

# Now, map the 'No impact' to 0, 'Positively' to 1, and 'Negatively' to -1 using impact_mapping
data['Impact of AI tools on job performance'] = data['Impact of AI tools on job performance'].map(impact_mapping)

# Convert values to numeric (coerce errors to NaN) for Section 23 & Section 20
data[section5_cols23] = data[section5_cols23].apply(pd.to_numeric, errors='coerce')
data[section5_cols20] = data[section5_cols20].apply(pd.to_numeric, errors='coerce')

# Fill missing values with overall mean of each column
data[section5_cols23] = data[section5_cols23].fillna(data[section5_cols23].mean())
data[section5_cols20] = data[section5_cols20].fillna(data[section5_cols20].mean())

# Round the values to the nearest integer
data[section5_cols23] = data[section5_cols23].round()
data[section5_cols20] = data[section5_cols20].round()

# Convert all columns to integer (nullable Int64 type)
data[section5_cols23] = data[section5_cols23].astype('Int64')
data[section5_cols20] = data[section5_cols20].astype('Int64')

# Save the cleaned Section 5 data to CSV
section5_cols = section5_cols23 + section5_cols20
data_section5 = data[section5_cols]
data_section5.to_csv('cleaned_section5.csv', index=False)
```

Final file base on all sections :

```
files = [
    'cleaned_section1.csv',
    'cleaned_section2.csv',
    'cleaned_section3.csv',
    'cleaned_section4.csv',
    'cleaned_section5.csv'
]

# Read each CSV file into a DataFrame
dfs = [pd.read_csv(file) for file in files]

# Concatenate all DataFrames horizontally (assuming rows align by index)
merged_data = pd.concat(dfs, axis=1)

# Export the merged DataFrame to a single CSV file
merged_data.to_csv('merged_cleaned_data.csv', index=False)
```