Part A

```
import pandas as pd
import numpy as np

df_teacher = pd.DataFrame({
    "name": ["Pep Guardiola", "Jurgen Klopp", "Mikel Arteta", "Zinadine
Zidane"],
    "married": [True, True, False, True],
    "school": ["Manchester High School", "Liverpool High School", "Arsenal
High", np.nan]
})
df_student = pd.DataFrame({
    "teacher": ["Mikel Arteta", "Mikel Arteta", "Pep Guardiola", "Jurgen
Klopp", "Jurgen Klopp", "Jurgen Klopp",
                "Pep Guardiola", "Pep Guardiola", "Mikel Arteta"],
    "name": ["Bukayo Saka", "Gabriel Martinelli", "Jack Grealish", "Roberto
Firmino", "Andrew Robertson",
             "Darwin Nunez", "Ederson Moraes", "Manuel Akanji", "Thomas
Partey"],
    "age": [21, 21, 27, 31, 28, 23, 29, 27, 29],
    "height": ['2.1m', '2.1m', '2.1m', '2.1m', '2.1m', '2.1m', '2.1m',
'2.1m',
               '2.1m'],

})

# Merge both dataframe into single dataframe
data = pd.merge(df_teacher, df_student, left_on='name', right_on='teacher')
# Change columns into meaningfull columns
data["student"] = data["name_y"]
data["teacher"] = data["name_x"]
# Drop the changed columns names
data.drop(columns=["name_x", "name_y"], axis=0, inplace=True)

# Define columns for each student in the list
students_columns = ["student", "age", "height"]

def formatData(group):
    formatedData = {
        'teacher': group['teacher'].iloc[0],
        'married': group['married'].iloc[0],
        'school': group['school'].iloc[0],
        'students': group[students_columns].to_dict(orient='records')
    }
    return formatedData

def teachersList():
    list = data.groupby('teacher').apply(lambda group:
formatData(group)).tolist()
    teachers_df = pd.DataFrame(list)
    return teachers_df


if __name__ == '__main__':
    print(teachersList().to_json(orient='records'))
```

PART B

```
 import pandas as pd
import numpy as np

df_teacher = pd.DataFrame({
    "name": ["Pep Guardiola", "Jurgen Klopp", "Mikel Arteta", "Zinadine
Zidane"],
    "married": [True, True, False, True],
    "school": ["Manchester High School", "Liverpool High School", "Arsenal
High", np.nan]
})
df_student = pd.DataFrame({
    "teacher": ["Mikel Arteta", "Mikel Arteta", "Pep Guardiola", "Jurgen
Klopp", "Jurgen Klopp", "Jurgen Klopp",
                "Pep Guardiola", "Pep Guardiola", "Mikel Arteta"],
    "name": ["Bukayo Saka", "Gabriel Martinelli", "Jack Grealish", "Roberto
Firmino", "Andrew Robertson",
             "Darwin Nunez", "Ederson Moraes", "Manuel Akanji", "Thomas
Partey"],
    "age": [21, 21, 27, 31, 28, 23, 29, 27, 29],
    "height": ['2.1m', '2.1m', '2.1m', '2.1m', '2.1m', '2.1m', '2.1m',
'2.1m',
               '2.1m'],
    "weight":
['80kg','70kg','690kg','73kg','60kg','70kg','80kg','88kg','74kg',]


})

# Merge both dataframe into single dataframe
data = pd.merge(df_teacher, df_student, left_on='name', right_on='teacher')
# Change columns into meaningfull columns
data["student"] = data["name_y"]
data["teacher"] = data["name_x"]
# Drop the changed columns names
data.drop(columns=["name_x", "name_y"], axis=0, inplace=True)

# Define columns for each student in the list
students_columns = ["student", "age", "height",'weight']

def formatData(group):
    formatedData = {
        'teacher': group['teacher'].iloc[0],
        'married': group['married'].iloc[0],
        'school': group['school'].iloc[0],
        'students': group[students_columns].to_dict(orient='records')
    }
    return formatedData

def teachersList():
    list = data.groupby('teacher').apply(lambda group:
```

```
formatData(group)).tolist()
    teachers_df = pd.DataFrame(list)
    return teachers_df


if __name__ == '__main__':
    print(teachersList().to_json(orient='records'))
```