PFA two ipynb files comprising questions based on pandas and the required data file. Write the Python statements corresponding to each and every question and submit the IPYNB files..

Write a Pandas program to select the rows the score is between 15 and 20 (inclusive). Sample DataFrame: exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'], 'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19], 'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1], 'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']} labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']

```
import pandas as pd
import numpy as np
exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matt
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
df = pd.DataFrame(exam_data)
print(df[df.score.between(15, 20)])
```

```
       name  score  attempts qualify
2  Katherine   16.5         2     yes
5    Michael   20.0         3     yes
9      Jonas   19.0         1     yes
```

Write a Pandas program to add a column named "column1" in the sixth position of the "coalpublic2013.xlsx" excel sheet and ll it with NaN values

```
import pandas as pd
import numpy as np
from google.colab import drive
drive.mount('/content/gdrive')
df = pd.read_excel('/content/gdrive/MyDrive/PANDAS Assignment/coalpublic2013.xlsx')
df.insert(5, "column1", np.nan)
print(df.columns)
```

```
Drive already mounted at /content/gdrive; to attempt to forcibly remount, call drive
Index(['Year', 'MSHA ID', 'Mine_Name', 'Production', 'Labor_Hours', 'column1'], dtype
```

```
import pandas as pd
import numpy as np
from google.colab import drive
df = pd.read_excel('/content/gdrive/MyDrive/PANDAS Assignment/coalpublic2013.xlsx')
print(df[df.Mine_Name.str.startswith('P')] )
```

```
    Year  MSHA ID                     Mine_Name Production Labor_Hours
13  2013   103332                Powhatan Mine   1,40,521       61394
18  2013   102976  Piney Woods Preparation Plant          0       14828
19  2013   102976  Piney Woods Preparation Plant          0       23193
46  2013   103321                Poplar Springs   1,89,370       76366
```

```
import pandas as pd
import numpy as np
from google.colab import drive
df = pd.read_csv('/content/gdrive/MyDrive/PANDAS Assignment/employee.csv')
df.hire_date=pd.to_datetime(df.hire_date)
print(df[df['hire_date'] >'2007/01/01'])
```

```
    emp_id first_name   last_name  hire_date
4      104      Bruce       Ernst 2007-05-21
7      107      Diana     Lorentz 2007-02-07
13     113       Luis        Popp 2007-12-07
19     119      Karen  Colmenares 2007-08-10
```

```
import pandas as pd
import numpy as np
df = pd.read_csv('/content/gdrive/MyDrive/PANDAS Assignment/Automobile_data.csv')
Average=df.groupby(by='company')
print(Average['average-mileage'].mean())
```

```
company
alfa-romero      20.333333
audi             20.000000
bmw              19.000000
chevrolet        41.000000
dodge            31.000000
honda            26.333333
isuzu            33.333333
jaguar           14.333333
mazda            28.000000
mercedes-benz    18.000000
mitsubishi       29.500000
nissan           31.400000
porsche          17.000000
toyota           28.714286
volkswagen       31.750000
volvo            23.000000
Name: average-mileage, dtype: float64
```

```
import pandas as pd
import numpy as np
df = pd.read_csv('/content/gdrive/MyDrive/PANDAS Assignment/Automobile_data.csv')
print(max(df.price))
```

```
45400.0
```