Author : Jashandeep Singh

Exploratory Data Analysis of the Energy Dataset

A1. Investigating the Energy Generation data for Victoria

1. First, read the data for Victoria state into a data frame. You will observe that some values for the fuel types (eg. Black coal etc.) are missing or have 'Nan'. To handle it, replace these values with zero (using appropriate python code) before proceeding with the rest of the questions.

Code:

```
# importing the libraries

import pandas as pd

import matplotlib.pyplot as plt

%matplotlib inline # for better plot visualisation

energy_data_df = pd.read_excel('energy_data.xlsx', sheet_name='VIC')

energy_data_df.fillna(0, inplace=True)  #changing null values as 0

energy_data_df #showing the df
```

Output:

	State	Fuel_Type	Category	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
0	VIC	Black coal	Non-renewable fuels	0.0	0.0	0.0	0.0	0.0	0.0	0.0	31.9	36.3	0.0
1	VIC	Brown coal	Non-renewable fuels	52094.0	51541.7	51066.5	52059.7	45317.6	43977.7	48336.8	46202.2	43557.8	36067.0
2	VIC	Natural gas	Non-renewable fuels	1451.9	1697.9	1289.8	1142.5	3247.7	3239.3	2390.9	1892.2	2658.7	3899.4
3	VIC	Oil products	Non-renewable fuels	20.4	1.1	38.2	4.0	10.7	145.8	156.0	70.6	109.3	164.8
4	VIC	Other	Non-renewable fuels	61.1	115.9	114.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	VIC	Biomass	Renewable fuels	292.0	303.4	339.6	859.3	845.1	886.9	672.2	747.7	694.5	661.8
6	VIC	Wind	Renewable fuels	573.0	1406.0	1434.4	1416.2	2005.1	2771.9	3067.8	3341.8	3560.9	4224.2
7	VIC	Hydro	Renewable fuels	557.8	843.7	1118.5	1047.4	940.3	1103.0	1170.9	1207.6	824.8	785.3
8	VIC	Large-scale solar PV	Renewable fuels	0.0	0.0	0.0	0.0	0.0	4.4	9.1	11.5	13.8	39.4
9	VIC	Small-scale solar PV	Renewable fuels	24.1	66.2	205.5	378.7	580.1	674.2	874.8	-1056.1	1231.7	1481.2
10	VIC	Geothermal	Renewable fuels	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Explanation: All the null values have been replaced as "0" in the dataset.

1.a. Using Python, plot the total energy generation in Victoria over the time period covered in the dataset (2009 to 2018). Describe the trend you see in the overall energy generation for the given time period.

Code:

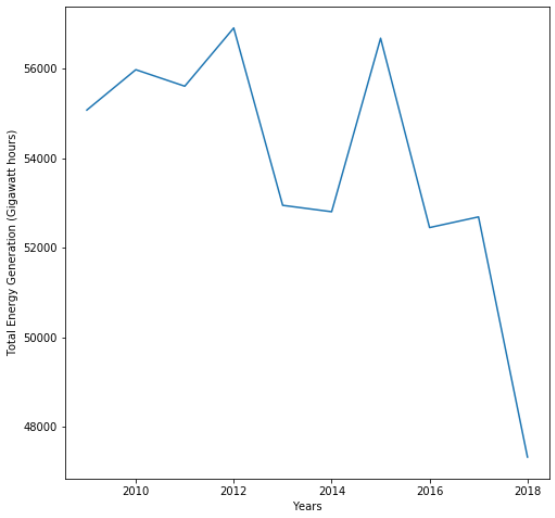
```
plt.ylabel("Total Energy Generation (Gigawatt hours)") #labeling the axis of the plot
plt.xlabel("Years")

energy_data_df.sum(axis=0 , numeric_only=True).plot(figsize=(8,8)) #Using the parameter numeric_only for filtering numerical values.

energy_data_df.sum(axis=0 , numeric_only=True #showing the output
```

Output:

1.

```
2009    55074.3
2010    55975.9
2011    55607.0
2012    56907.8
2013    52946.6
2014    52803.2
2015    56678.5
2016    52449.4
2017    52687.8
2018    47323.1
dtype: float64
```

2.



Explanation:

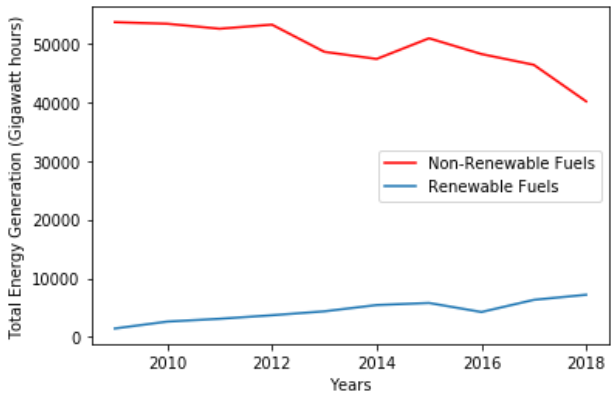
From the output and the plot, it is observed that the year 2012 had the highest Total Energy Generation and the year 2018 marked the lowest Total Energy Generation. Major fluctuations were also observed between years 2012-2013 (decreasing), 2014-2016 (increasing and decreasing) and, 2017-2018 (decreasing). Overall trend displayed by the plot states that Total Energy Generation in Victoria has decreased.

1.b. Draw a new plot showing the trend in total renewable and non-renewable energy generation for the same time period? What trend can you observe from this graph?

Code:

```
non_renewable_sum = energy_data_df.loc[energy_data_df['Category'] == 'Non-renewable fuels'].sum(axis=0 , numeric_only=True)
renewable_sum = energy_data_df.loc[energy_data_df['Category'] == 'Renewable fuels'].sum(axis=0 , numeric_only=True)

plt.ylabel("Total Energy Generation (Gigawatt hours)") #labelling the axis of the plot
plt.xlabel("Years")

plt.plot(non_renewable_sum,  'r', label="Non-Renewable Fuels") # defining red colour to the plot and declaring labels for legend
plt.plot(renewable_sum, label="Renewable Fuels")
plt.legend() #showing the legend in the plot
```

Output:



Explanation:

It is observed that the usage of Non-Renewable fuels remained as the main source of fuel for the total energy generation as denoted by the y-axis of the plot, whereas renewable fuels were very less used in the energy generation. It is also

observed that the use of Renewable fuels as the source of energy generation is increasing with the years and use of Non-renewable fuels is decreasing with the years.

1.c. Draw a bar chart showing the breakdown of the different fuel types used for energy generation in 2009 vs in 2018? Explain your observation.

Code:

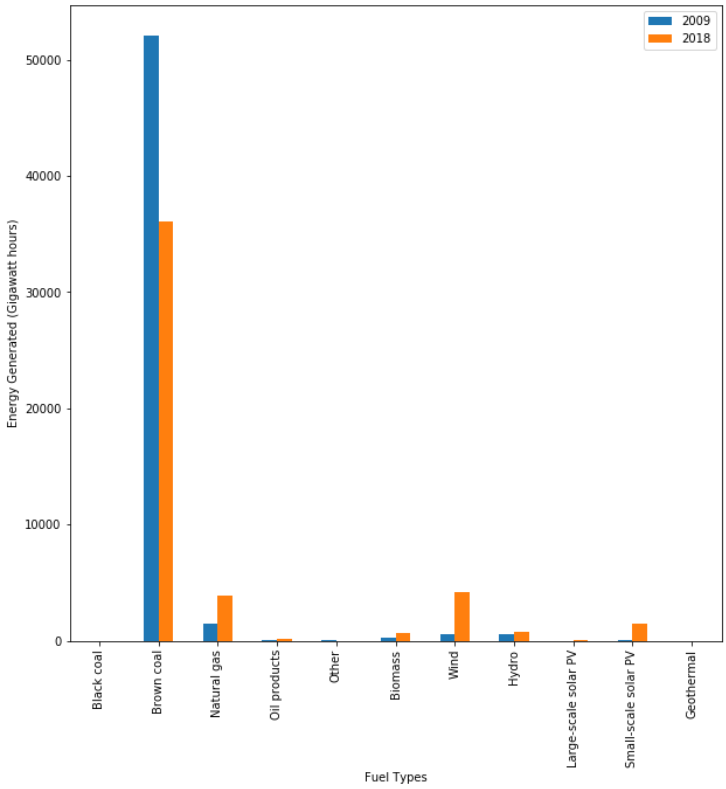
```
energy_data_df[["Fuel_Type", 2009, 2018]].set_index('Fuel_Type').plot.bar(figsize=(10,10))
plt.xlabel("Fuel Types")
plt.ylabel("Energy Generated (Gigawatt hours)")

energy_data_df[["Fuel_Type", 2009, 2018]].set_index('Fuel_Type') #showing the df
```

Output:

1.

Fuel_Type	2009	2018
Black coal	0.0	0.0
Brown coal	52094.0	36067.0
Natural gas	1451.9	3899.4
Oil products	20.4	164.8
Other	61.1	0.0
Biomass	292.0	661.8
Wind	573.0	4224.2
Hydro	557.8	785.3
Large-scale solar PV	0.0	39.4
Small-scale solar PV	24.1	1481.2
Geothermal	0.0	0.0

2.



Explanation:

It is observed from the output that Brown Coal has been the major source of Energy Generation in both 2009 and 2018, where the usage is decreased in 2018 as compared to 2009. Also, no energy was generated from Black Coal and Geothermal fuel types in both the years. It is also observed that in Victoria, energy generation from Other fuel type has been stopped in 2018 as compared to 2009 and that Victoria has also started generating energy from Large-scale solar PV in the year 2018 as compared to 2009.

1.d. What was the most used energy resource (fuel-type) in 2015? Which renewable fuel type was the least used in 2015?

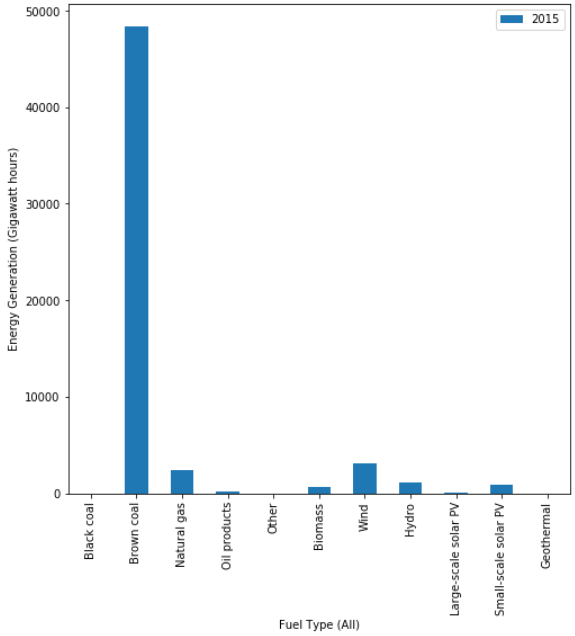
Code:

```
energy_data_df[["Fuel_Type",2015]].set_index('Fuel_Type').plot.bar(figsize=(8,8)) #Selecting index for plotting fuel type and year

plt.xlabel("Fuel Type (All)")
plt.ylabel("Energy Generation (Gigawatt hours)")

x = energy_data_df[energy_data_df.Category == "Renewable fuels"] #defining new dataframe with only Renewable fuels

x[["Fuel_Type",2015]].set_index('Fuel_Type').plot.bar(figsize=(8,8))
```

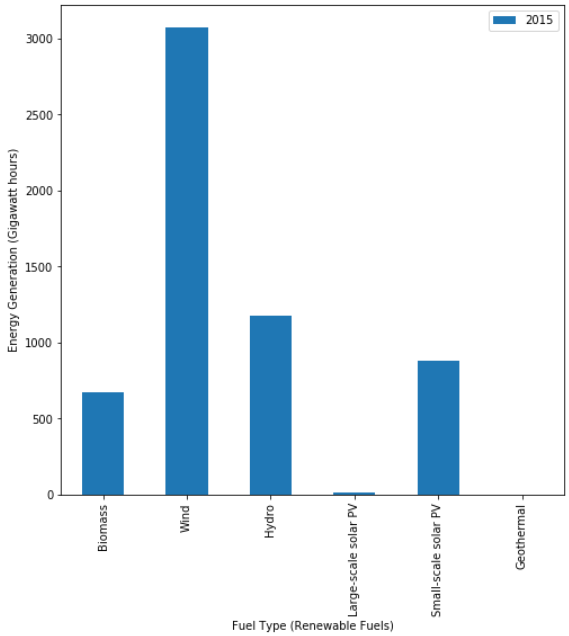
```
plt.xlabel("Fuel Type (Renewable Fuels)")
plt.ylabel("Energy Generation (Gigawatt hours)")
```

Output:

1. Energy Generation v/s Fuel Type(All) plot:



2. Energy Generation v/s Fuel Type (Renewable Fuels) plot:



Explanation:

From the output it is observed that, "Brown Coal" was the most used energy resource in the year 2015 as observed by the Energy Generation v/s Fuel Type(All) plot and the renewable fuel type "Geothermal" was the least used in the year 2015 as observed by the Energy Generation v/s Fuel Type (Renewable Fuels) plot.

1.e. Draw a plot showing the percentage of Victoria's energy generation coming from Renewable vs Non-Renewable energy sources over the period 2009 to 2018. What can you say about the trend you observe?

Code:

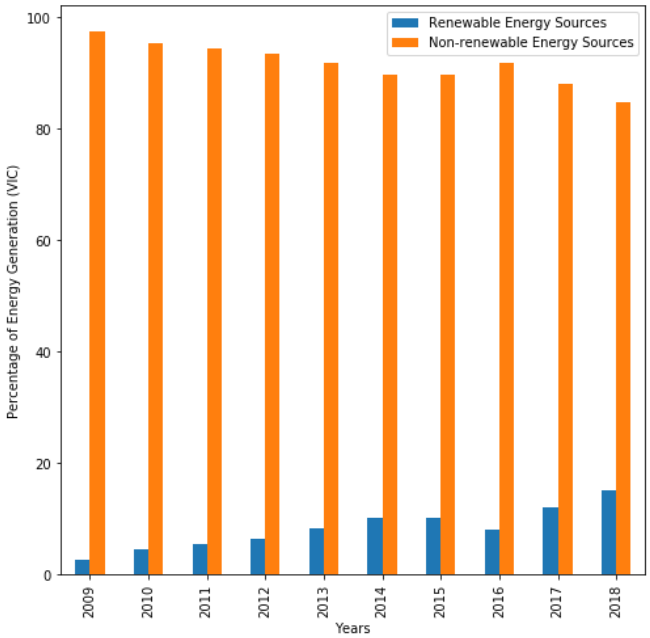
```
total_energy = non_renewable_sum + renewable_sum    # variables from part A1.1.b. , total_energy defines the total energy produced in that year

df_energy_generation = pd.DataFrame({
                        "Renewable Energy Sources": (renewable_sum/total_energy)*100,
                        "Non-renewable Energy Sources": (non_renewable_sum/total_energy)*100,
                       })    # years are already defined in index for plotting


df_energy_generation.plot.bar(figsize=(8,8)) #changing the size of the plot

plt.xlabel("Years")
plt.ylabel("Percentage of Energy Generation (VIC)")

df_energy_generation # showing the df
```

Output:

1.

	Renewable Energy Sources	Non-renewable Energy Sources
2009	2.627178	97.372822
2010	4.679335	95.320665
2011	5.571241	94.428759
2012	6.504556	93.495444
2013	8.254732	91.745268
2014	10.303163	89.696837
2015	10.223983	89.776017
2016	8.107814	91.892186
2017	12.006005	87.993995
2018	15.197441	84.802559

2.



Explanation:

It is observed from the output that from the years 2009 to 2018, the percentage of Victoria's energy coming from Renewable Energy Sources is increasing. Similarly, it is also observed that the percentage of Victoria's Energy coming from Non-renewable energy sources is decreasing. Furthermore, from the plot, slight fluctuations in the trend were also observed in the year 2016, where energy generation from renewable sources was decreased and from Non-renewable sources was increased.

1.f. Using a linear regression model, predict what percentage of Victoria's energy generation will come from Renewable energy sources in the year 2030, 2100? Do the predictions seem reasonable?

Code:

```
percentage_renewable_sum  = (renewable_sum/total_energy)*100   # variables from A1.1.e

# converting series to dataframe
df_percentage = percentage_renewable_sum.to_frame().reset_index()


df_percentage.columns = ["Years", "Percentage"]

df_percentage.set_index("Years")



plt.scatter(df_percentage["Years"] ,df_percentage["Percentage"] )
plt.ylabel("Percentage")
plt.xlabel("Years")



# linear regression model

from scipy import stats

from scipy.stats import linregress


# linregress() will help us determine slope and intercept of the function.

slope, intercept, r_value, p_value, std_err = linregress(df_percentage["Years"] ,df_percentage["Percentage"] )

print("Slope is %f" %slope , "Intercept is %f" %intercept)

line = [slope*xi + intercept for xi in df_percentage["Years"]]   # y = mx+c, determining the regression line on current data

plt.plot(df_percentage["Years"],line,'r-', linewidth=3 , label="Regression line")
plt.scatter(df_percentage["Years"] ,df_percentage["Percentage"] ,label="Original data" )
plt.legend()
plt.show()

# generating prediction values by using y = mx+c, where x is Years, m is the slope, c is the intercept.

# for year 2030
x_predict = 2030
y_predict = slope * x_predict + intercept

print("The prediction for year 2030 is: %f percent"  %y_predict)

#for year 2100
x_predict = 2100
y_predict = slope * x_predict + intercept
print("The prediction for year 2100 is: %f percent"  %y_predict)

df_percentage # showing the df
```

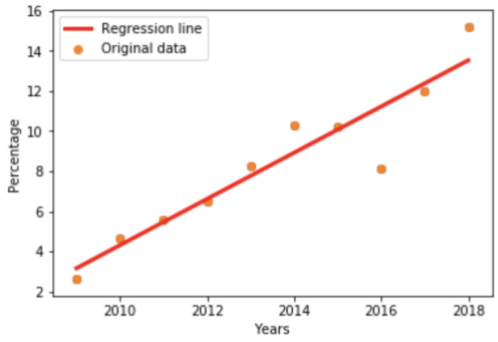
Output:

1.

	Years	Percentage
0	2009	2.627178
1	2010	4.679335
2	2011	5.571241
3	2012	6.504556
4	2013	8.254732
5	2014	10.303163
6	2015	10.223983
7	2016	8.107814
8	2017	12.006005
9	2018	15.197441

2.

```
Slope is 1.153386 Intercept is -2313.994407
```




```
The prediction for year 2030 is: 27.378408 percent
The prediction for year 2100 is: 108.115401 percent
```

Explanation:

It can be observed from the output that by the trend that Victoria's energy generation by Renewable fuels is progressing as the year's progress. From the model, it is predicted that in the year 2030, that is after 22 years from the given data, 27.378408 % of Victoria's Energy Generation will come from Renewable Fuels, which is promising given the shift of usage from non-renewable to renewable fuels that was observed from the data. In the year 2100, that is 82 years after the given data, it is predicted that 108.115401 % of Victoria's Energy Generation will come from Renewable Fuels, which interprets as Victoria will completely generate its energy from Renewable Fuels. For the year 2030, the prediction is quite reasonable, as observed by the trend, but for the year 2100, the predicted value might not be reasonable since our data set for the model is quite small for the prediction of the value that is 82 years ahead in future.

A2. Investigating the Energy Generation data for Australia.

1. Let's do some further investigation by combining the data for all the states and territories in Australia. Read the data for the rest of the states and merge them in a single data frame. (Hint: you can use a combination of merge, melt or concat operators to get your data in a format suitable for answering the following questions)

Code:

```
all_energy_data_df = pd.concat(pd.read_excel("energy_data.xlsx", sheet_name=None), ignore_index=True)

all_energy_data_df.fillna(0, inplace=True)  #changing null values as 0, for easier auditing

all_energy_data_df # showing the df
```

Output:

	State	Fuel_Type	Category	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
0	VIC	Black coal	Non-renewable fuels	0.0	0.0	0.0	0.0	0.0	0.0	0.0	31.9	36.3	0.0
1	VIC	Brown coal	Non-renewable fuels	52094.0	51541.7	51066.5	52059.7	45317.6	43977.7	48336.8	46202.2	43557.8	36067.0
2	VIC	Natural gas	Non-renewable fuels	1451.9	1697.9	1289.8	1142.5	3247.7	3239.3	2390.9	1892.2	2658.7	3899.4
3	VIC	Oil products	Non-renewable fuels	20.4	1.1	38.2	4.0	10.7	145.8	156.0	70.6	109.3	164.8
4	VIC	Other	Non-renewable fuels	61.1	115.9	114.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...
72	WA	Wind	Renewable fuels	675.1	664.0	719.3	1278.7	1300.2	1578.8	1643.2	1476.5	1616.7	1593.0
73	WA	Hydro	Renewable fuels	0.0	0.0	0.0	0.0	221.3	205.4	206.1	217.0	212.6	217.9
74	WA	Large-scale solar PV	Renewable fuels	0.0	0.0	0.0	0.0	0.0	24.1	23.9	28.3	34.3	45.6
75	WA	Small-scale solar PV	Renewable fuels	18.7	57.5	198.1	331.4	449.5	495.8	622.1	760.8	941.1	1196.4
76	WA	Geothermal	Renewable fuels	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

77 rows × 13 columns

1.a. Plot a column chart showing the total energy generated in Australia by fuel type in the year 2018.

Code:

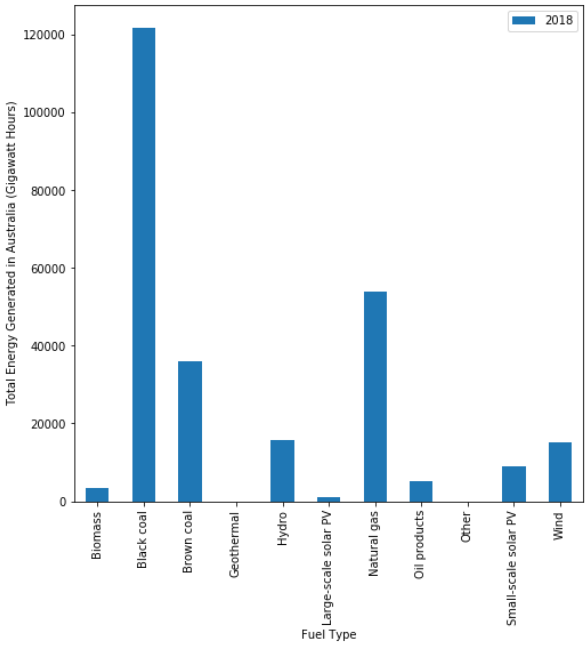
```
f = all_energy_data_df[["Fuel_Type", 2018]]
f.groupby("Fuel_Type").sum().plot.bar(figsize = (8,8)) # Total energy generated
plt.xlabel("Fuel Type")
plt.ylabel("Total Energy Generated in Australia (Gigawatt Hours)")
f.groupby("Fuel_Type").sum() # showing the data
```

Output:

1.

Fuel_Type	2018
Biomass	3534.1
Black coal	121592.6
Brown coal	36067.0
Geothermal	0.0
Hydro	15838.9
Large-scale solar PV	1007.6
Natural gas	53839.4
Oil products	5262.6
Other	0.0
Small-scale solar PV	8922.4
Wind	14989.5

2.



Explanation:

The output provides the total energy generated by each fuel type in Australia in the year of 2018. It can be observed that Black Coal was the biggest source of energy in Australia for that year.

1.b. Which state had the highest energy production in 2018? What is the ratio (percentage breakdown) of renewable vs non-renewable energy production for that state in 2018?

Code:

```
g = all_energy_data_df[["State", "Fuel_Type", 2018]]

total_state_energy = g.groupby("State").sum() #State is the index of the df

max_state = total_state_energy[[2018]].idxmax() #returns the Name for the state with highest energy production
print(max_state)

df_NSW = all_energy_data_df.loc[all_energy_data_df["State"]== "NSW"].reset_index()
df_NSW_energy = df_NSW[["Fuel_Type", "Category", 2018]]
df_NSW_energy # df of NSW for year 2018

NSW_renewable_sum = df_NSW_energy.loc[df_NSW_energy["Category"] == "Renewable fuels"].sum(axis=0 , numeric_only=True)
NSW_non_renewable_sum = df_NSW_energy.loc[df_NSW_energy["Category"] == "Non-renewable fuels"].sum(axis=0 , numeric_only=True)
NSW_total_sum = NSW_renewable_sum + NSW_non_renewable_sum

NSW_energy_breakdown = pd.DataFrame({ "Renewable Energy Production Percentage" : (NSW_renewable_sum/NSW_total_sum)*100,
                    "Non-renewable Energy Production Percentage"
:(NSW_non_renewable_sum/NSW_total_sum)*100
                    }) # the dataframe consists of the percentage values
```

```
NSW_energy_breakdown.plot.bar(figsize = (8,8))
plt.xlabel("Year")
plt.ylabel("Percentage Energy Production")
```

```
NSW_energy_breakdown # showing the df
```

Output:

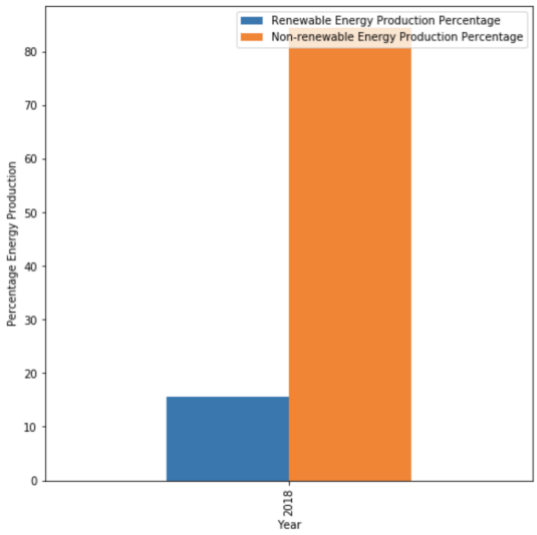
1.

```
2018    NSW
dtype: object
```

2.

	Renewable Energy Production Percentage	Non-renewable Energy Production Percentage
2018	15.660312	84.339688

3.



Explanation:

From the output, it is observed that the state with the highest energy production in 2018 is NSW. In the State of NSW, in the year 2018, 15.660312 % of total energy generation came from Renewable Fuels and 84.339688 % of total energy generation came from Non-renewable Fuels.

1.c. Draw a plot showing the percentage of energy generation from renewable energy sources for each state over the period 2009 to 2018. From your graph, which state do you think is making the most progress towards adopting green energy? Provide a reason for your answer.

Code:

```
# df with total energy generation
total_energy = all_energy_data_df.groupby("State").sum()

# new df with total renewable fuel energy generation
x = all_energy_data_df.groupby(["State", "Category"]).sum()
y = x.reset_index() # to reset the index from State and Category
p = y.loc[y.Category == "Renewable fuels"] #filtering the Renewable fuels
# formatting and setting index as "State"
del p["Category"]
total_energy_renewable = p.set_index("State")
```

 # Calculating the percentage of energy generation from renewable energy sources for each state over the period 2009 to 2018

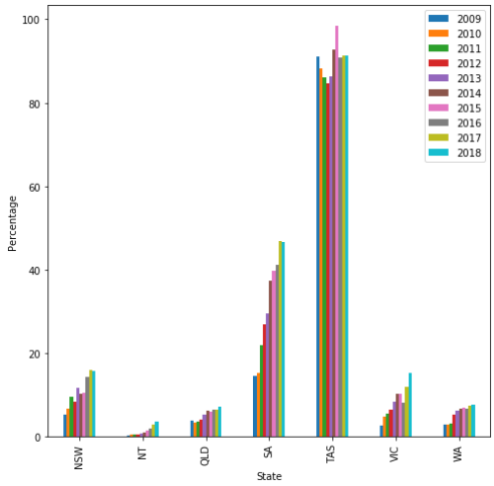
```
l = total_energy_renewable.div(total_energy)*100
l.plot.bar(figsize = (8,8))
plt.xlabel("State")
plt.ylabel("Percentage")

l # showing the df
```

Output:

1.

State	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
NSW	5.278109	6.771295	9.536412	8.447289	11.630281	10.177022	10.574166	14.231296	16.037979	15.660312
NT	0.331919	0.413274	0.462648	0.552802	0.706010	0.963675	1.478145	1.911347	2.921807	3.513067
QLD	3.884591	3.476066	3.606199	4.140304	5.207971	6.292163	6.044747	6.463106	6.497393	7.151522
SA	14.424195	15.350651	21.855236	26.845457	29.485390	37.472847	39.854294	41.250526	46.962737	46.633725
TAS	91.045922	88.180493	86.172198	84.712464	86.347701	92.823264	98.443535	90.832721	91.329115	91.312427
VIC	2.627178	4.679335	5.571241	6.504556	8.254732	10.303163	10.223983	8.107814	12.006005	15.197441
WA	2.933825	2.920728	3.221948	5.373827	6.133248	6.619774	6.940360	6.653088	7.462474	7.744543

2.



Explanation:

It is observed that the state "SA" is making the most progress in adopting green energy because the graph shown by the data set of State SA shows the highest positive slope (the inclination of the data points plotted on the plot), that can also be explained as, the difference between the data points marked at year 2009 and year 2018, for state SA, is the highest. To support this we take reference from the df output.

The difference between 2009 and 2018 values calculated for the state SA is 32.20953, which is the highest value as compared to the difference between 2009 and 2018 values of other states. The higher difference value tells us that the state has made the highest progress towards adopting green energy.

A3. Visualising the Relationship over Time

Now let's look at the relationship between all variables impacting the energy generation over time. Ensure that you have combined all the data from the different states. Ensure that your data is aggregated by year, state, the total energy produced (total_production), and has a separate column for each of the fuel types.

1. Use Python to build a Motion Chart, that visualises the energy production trend for Australia over time. The motion chart should show the units of energy production using Wind on the x-axis, the energy production using Natural gas on the y-axis, the colour represents the states/territories the bubble size should show the total_production.

Code:

```
from motionchart.motionchart import MotionChart # importing for motionchart
d = all_energy_data_df.drop(["Category"], axis=1)
d_1 = d.melt(id_vars = ["State", "Fuel_Type"], var_name = ["Years"], value_name = "Year_Value")

d_1.set_index("Years", "State")
d_chart= d_1.pivot_table("Year_Value", ["Years", "State"], "Fuel_Type").reset_index().set_index("Years", "State")

d_chart.loc[:, "total_production"] = d_chart.sum(axis=1, numeric_only = True)
#https://stackoverflow.com/questions/53414960/how-do-i-create-a-sum-row-and-sum-column-in-pandas. , retrieved 28th April, 2020

d_chart.reset_index(inplace=True)
d_chart.columns.name = None

d_chart # showing the df


# html code block for clean visual output of motion chart
%%html
<style>
.output_wrapper, .output {
    height:auto !important;
    max-height:1000px;  /* your desired max-height here */
}
.output_scroll {
    box-shadow:none !important;
    webkit-box-shadow:none !important;
}
</style>


# defining variables of motionchart
```

```
mChart = MotionChart(df = d_chart, key='Years', x='Wind', y='Natural gas', xscale='linear', yscale='linear',
                     size='total_production', color='State', category='State')

mChart.to_notebook() # motion chart display
```

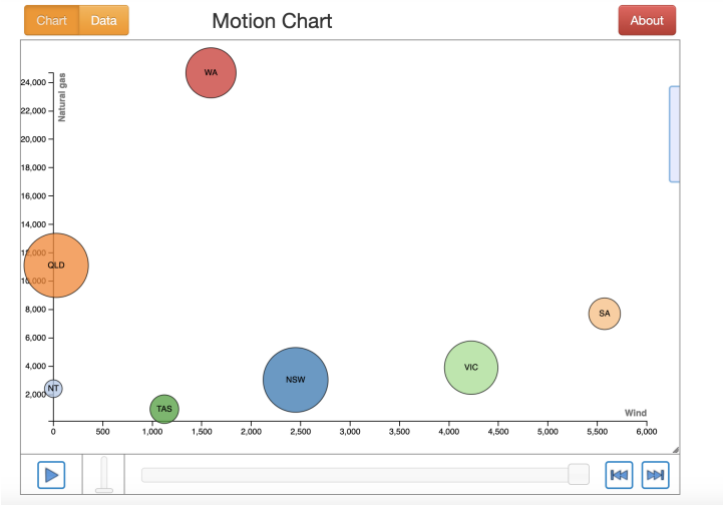
Output:

1.

	Years	State	Biomass	Black coal	Brown coal	Geothermal	Hydro	Large-scale solar PV	Natural gas	Oil products	Other	Small-scale solar PV	Wind	total_production
0	2009	NSW	674.7	67650.3	0.0	0.0	3173.7	0.0	2445.1	24.7	336.6	36.3	41.3	74382.7
1	2009	NT	9.2	0.0	0.0	0.0	0.0	0.0	1873.4	1039.3	0.0	0.5	0.0	2922.4
2	2009	QLD	1588.5	50882.2	0.0	0.5	820.2	0.0	9257.2	711.4	691.8	50.8	27.3	64029.9
3	2009	SA	86.1	0.0	4887.4	0.0	4.1	0.0	7371.4	63.2	337.4	23.4	2020.2	14793.2
4	2009	TAS	25.1	0.0	0.0	0.0	7313.6	0.0	705.3	64.5	0.0	1.8	486.9	8597.2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
65	2018	QLD	1434.1	53480.2	0.0	0.0	657.3	171.5	11112.1	1001.9	0.0	2760.9	28.5	70646.5
66	2018	SA	89.1	0.0	0.0	0.0	6.4	20.0	7703.5	137.5	0.0	1163.4	5572.9	14692.8
67	2018	TAS	24.6	0.0	0.0	0.0	9327.4	0.6	976.8	33.0	0.0	138.9	1122.2	11623.5
68	2018	VIC	661.8	0.0	36067.0	0.0	785.3	39.4	3899.4	164.8	0.0	1481.2	4224.2	47323.1
69	2018	WA	153.0	10960.9	0.0	0.0	217.9	45.6	24687.9	2540.9	0.0	1196.4	1593.0	41395.6

70 rows × 14 columns

2.



2. Run the visualisation from start to end. (Hint: In Python, to speed up the animation, set the timer bar next to the play/pause button to the minimum value.) And then answer the following questions:

2.a. Comment generally on the trend you see on reliance on wind energy vs reliance on natural gas for each Australian state overtime. Is it logical to say if there is a relationship between the two variables?

Explanation:

The trend on reliance shown by each Australian state is as follows:

1. NT shows the least reliance on wind in the time frame and is mostly dependent upon natural gas throughout.
2. QLD shows the most reliance on Natural Gas and very less increase in reliance to wind in the time frame. It also develops a decrease in reliance to Natural gas from years 2015 onwards.
3. TAS develops a significant reliance on wind energy in the years of 2014 and 2015 and maintains the same.
4. WA overall shows a linear trend of increase of reliance on both natural gas and wind, although for the period of 2011 to 2013 WA develops heavy reliance on wind.
5. NSW develops a linear trend of growth with both natural gas and wind but relies heavily on wind energy 2013 onwards.
6. VIC intially has a constant incresing reliance on wind but after 2012 also shows reliance on natural gas. Overall VIC shows the most reliance on Wind.
7. SA shows a constant reliance with Natural Gas but develops a strong reliance with wind in the time frame.

As observed with the Motion Chart, except state "NT", all the other states shows relationship between both the variables. In the time frame, there is increase and decrease of data as displayed by the bubbles.
This variation is only happening because of the relationship between both the variables Natural Gas and Wind (y-axis and x-axis). So logically for the motion chart, there is a relationship between the variables.

2.b. Which state relied most on natural gas for energy production in 2013? Please support your answer with any relevant python code and the motion chart screenshot.

Code:

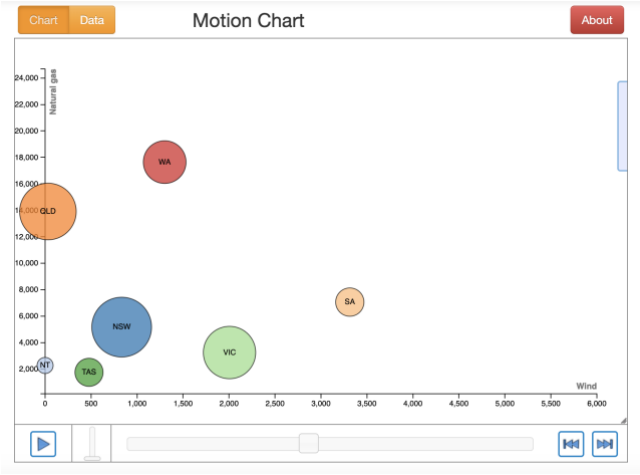
```
d_chart[["Years","State","Natural gas"]].loc[d_chart.Years == 2013].max() #max value for year 2013
```

Output:

1.

```
Years          2013
State            WA
Natural gas  17645.1
```

2.



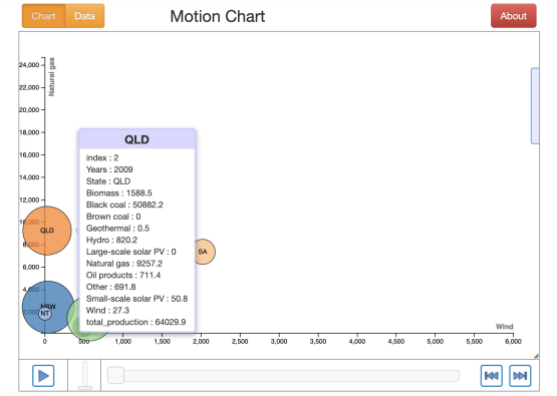
Explanation:

In the year 2013, the state “WA” had the highest reliance on Natural Gas, with the total energy production of 17645.1. It can also be observed from the motion chart as the bubble “WA” has the highest value over y-axis (Natural Gas) in the year 2013.

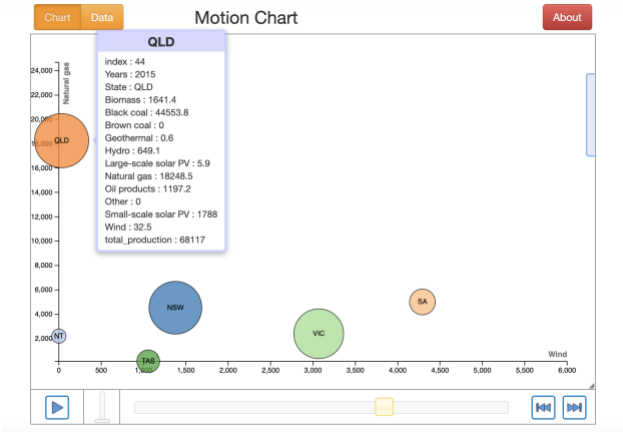
2.c. Comment on Queensland (QLD) states reliance trend on Natural gas between 2009 to 2018? What could be the reason contributing to this?

Screenshot of the Motion Chart:

Year 2009:



Year 2015:



The trend shown by the state “QLD” on the motion chart denotes that the state kept significant reliance on natural gas throughout the timeframe with having the maximum value for natural gas in the year 2015 and the minimum in the year 2009. The state relied very little on wind as a source of energy generation throughout the time frame.