```
In [1]: #Reading the Dataset
        #import libraries
        import pandas as pd
        import matplotlib.pyplot as plt
        import numpy as np
        import seaborn as sns
        data = pd.read_excel(r'D:\food_prices_dza.xlsx')
        data.head()
```

Out[1]:

	date	country	admin1	admin2	admin3	latitude	longitude	category	commodity	unit	price
0	2015-04-15	Algeria	Alger	NaN	Hydra	36.752887	3.042048	cereals and tubers	Bread	Unit	a
1	2015-04-15	Algeria	Alger	NaN	Hydra	36.752887	3.042048	cereals and tubers	Pasta	500 G	a
2	2015-04-15	Algeria	Alger	NaN	Hydra	36.752887	3.042048	cereals and tubers	Potatoes	KG	a
3	2015-04-15	Algeria	Alger	NaN	Hydra	36.752887	3.042048	cereals and tubers	Rice	KG	a
4	2015-04-15	Algeria	Alger	NaN	Hydra	36.752887	3.042048	meat, fish and eggs	Eggs	30 pcs	a

```
In [2]: #print the dataframe
        df = pd.DataFrame(data)
        df
```

Out[2]:

	date	country	admin1	admin2	admin3	latitude	longitude	category	commodity	u
0	2015-04-15	Algeria	Alger	NaN	Hydra	36.752887	3.042048	cereals and tubers	Bread	U
1	2015-04-15	Algeria	Alger	NaN	Hydra	36.752887	3.042048	cereals and tubers	Pasta	50
2	2015-04-15	Algeria	Alger	NaN	Hydra	36.752887	3.042048	cereals and tubers	Potatoes	
3	2015-04-15	Algeria	Alger	NaN	Hydra	36.752887	3.042048	cereals and tubers	Rice	
4	2015-04-15	Algeria	Alger	NaN	Hydra	36.752887	3.042048	meat, fish and eggs	Eggs	30
...	...	...	...	...	...	...	...	...	...	
1488	2021-05-15	Algeria	Tindouf	NaN	Tindouf	27.676101	-8.127653	miscellaneous food	Tea (sahm)	Pac
1489	2021-05-15	Algeria	Tindouf	NaN	Tindouf	27.676101	-8.127653	pulses and nuts	Beans (white)	
1490	2021-05-15	Algeria	Tindouf	NaN	Tindouf	27.676101	-8.127653	pulses and nuts	Lentils	
1491	2021-05-15	Algeria	Tindouf	NaN	Tindouf	27.676101	-8.127653	vegetables and fruits	Apples	
1492	2021-05-15	Algeria	Tindouf	NaN	Tindouf	27.676101	-8.127653	vegetables and fruits	Tomatoes	

1493 rows × 16 columns

Loading [MathJax]/extensions/Safe.js

```
In [3]:    #Dimensionality
           df.shape
```

```
Out[3]:    (1493, 16)
```

```
In [4]:    #Data type of each column
           df.dtypes
```
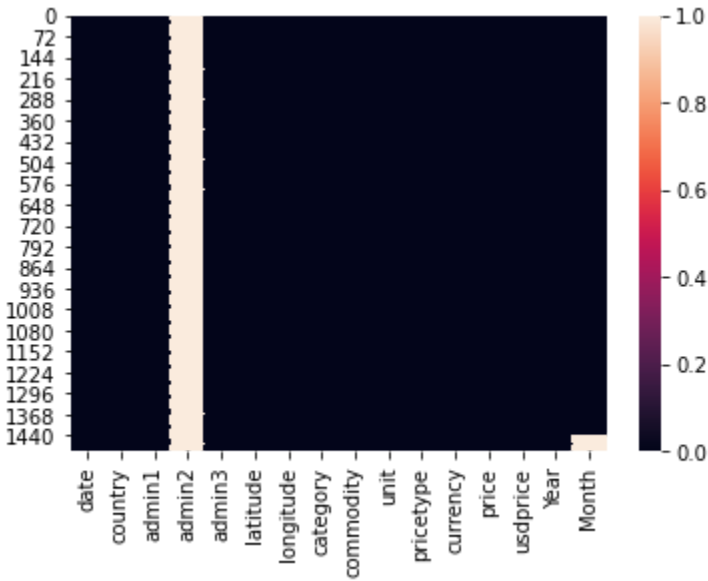
```
Out[4]:    date          object
           country       object
           admin1        object
           admin2        float64
           admin3        object
           latitude      float64
           longitude     float64
           category      object
           commodity     object
           unit          object
           pricetype     object
           currency      object
           price         float64
           usdprice      float64
           Year            int64
           Month         float64
           dtype: object
```

```
In [5]:    #data exploration and preparation
           #Question1
           sns.heatmap(df.isna())
```

```
Out[5]:    <AxesSubplot:>
```



```
In [6]:    dataframe = df.drop(['admin2'],axis=1)
           dataframe
```

Out[6]:

	date	country	admin1	admin3	latitude	longitude	category	commodity	unit	price
0	2015-04-15	Algeria	Alger	Hydra	36.752887	3.042048	cereals and tubers	Bread	Unit	a
1	2015-04-15	Algeria	Alger	Hydra	36.752887	3.042048	cereals and tubers	Pasta	500 G	a
2	2015-04-15	Algeria	Alger	Hydra	36.752887	3.042048	cereals and tubers	Potatoes	KG	a
3	2015-	Algeria	Alger	Hydra	36.752887	3.042048	cereals and tubers	Rice	KG	a

Loading [MathJax]/extensions/Safe.js

	date	country	admin1	admin3	latitude	longitude	category	commodity	unit	price
4	2015-04-15	Algeria	Alger	Hydra	36.752887	3.042048	meat, fish and eggs	Eggs	30 pcs	a
...	...	...	...	...	...	...	...	...	...	
1488	2021-05-15	Algeria	Tindouf	Tindouf	27.676101	-8.127653	miscellaneous food	Tea (sahm)	Packet	a
1489	2021-05-15	Algeria	Tindouf	Tindouf	27.676101	-8.127653	pulses and nuts	Beans (white)	KG	a
1490	2021-05-15	Algeria	Tindouf	Tindouf	27.676101	-8.127653	pulses and nuts	Lentils	KG	a
1491	2021-05-15	Algeria	Tindouf	Tindouf	27.676101	-8.127653	vegetables and fruits	Apples	KG	a
1492	2021-05-15	Algeria	Tindouf	Tindouf	27.676101	-8.127653	vegetables and fruits	Tomatoes	KG	a

1493 rows × 15 columns

```
In [7]:  dataframe = dataframe.drop(['country', 'pricetype', 'currency'],axis=1)
         dataframe
```

Out[7]:

	date	admin1	admin3	latitude	longitude	category	commodity	unit	price	usdpric
0	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Bread	Unit	10.0	0.0
1	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Pasta	500 G	60.0	0.4
2	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Potatoes	KG	80.0	0.5
3	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Rice	KG	90.0	0.6
4	2015-04-15	Alger	Hydra	36.752887	3.042048	meat, fish and eggs	Eggs	30 pcs	300.0	2.1
...	...	...	...	...	...	...	...	...	...	
1488	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	miscellaneous food	Tea (sahm)	Packet	200.0	1.4
1489	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	pulses and nuts	Beans (white)	KG	270.0	1.9
1490	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	pulses and nuts	Lentils	KG	185.0	1.3
1491	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	vegetables and fruits	Apples	KG	450.0	3.2
1492	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	vegetables and fruits	Tomatoes	KG	120.0	0.8

1493 rows × 12 columns

```
In [8]:  #true for nan values
         dataframe.isnull()
```

Out[8]:

	date	admin1	admin3	latitude	longitude	category	commodity	unit	price	usdprice	Year
0	False	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False	False

Loading [MathJax]/extensions/Safe.js

	date	admin1	admin3	latitude	longitude	category	commodity	unit	price	usdprice	Year
2	False	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...	...	...
1488	False	False	False	False	False	False	False	False	False	False	False
1489	False	False	False	False	False	False	False	False	False	False	False
1490	False	False	False	False	False	False	False	False	False	False	False
1491	False	False	False	False	False	False	False	False	False	False	False
1492	False	False	False	False	False	False	False	False	False	False	False

1493 rows × 12 columns

```
In [9]:  dataframe['Month'].fillna(dataframe['date'], inplace=True)
         dataframe
```

Out[9]:

	date	admin1	admin3	latitude	longitude	category	commodity	unit	price	usdpri
0	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Bread	Unit	10.0	0.0
1	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Pasta	500 G	60.0	0.4
2	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Potatoes	KG	80.0	0.5
3	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Rice	KG	90.0	0.6
4	2015-04-15	Alger	Hydra	36.752887	3.042048	meat, fish and eggs	Eggs	30 pcs	300.0	2.1
...	...	...	...	...	...	...	...	...	...	
1488	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	miscellaneous food	Tea (sahm)	Packet	200.0	1.4
1489	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	pulses and nuts	Beans (white)	KG	270.0	1.9
1490	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	pulses and nuts	Lentils	KG	185.0	1.3
1491	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	vegetables and fruits	Apples	KG	450.0	3.2
1492	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	vegetables and fruits	Tomatoes	KG	120.0	0.8

1493 rows × 12 columns

```
In [10]:  dataframe.at[635,'date'] = '2015-12-15'
          print(dataframe.at[635,'date'])

          dataframe.at[636,'date'] = '2016-1-15'
          print(dataframe.at[636,'date'])

          dataframe.at[637,'date'] = '2016-1-15'
          print(dataframe.at[637,'date'])

          2015-12-15
```

2016-1-15

In [11]:
```python
#Question2
dataframe = dataframe.rename(columns = {'admin1': 'State','admin3': 'City','price': 'DzdP
dataframe
```

Out[11]:

	date	State	City	latitude	longitude	category	commodity	unit	DzdPrice	usdp
0	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Bread	Unit	10.0	
1	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Pasta	500 G	60.0	
2	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Potatoes	KG	80.0	
3	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Rice	KG	90.0	
4	2015-04-15	Alger	Hydra	36.752887	3.042048	meat, fish and eggs	Eggs	30 pcs	300.0	
...	...	...	...	...	...	...	...	...	...	
1488	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	miscellaneous food	Tea (sahm)	Packet	200.0	
1489	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	pulses and nuts	Beans (white)	KG	270.0	
1490	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	pulses and nuts	Lentils	KG	185.0	
1491	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	vegetables and fruits	Apples	KG	450.0	
1492	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	vegetables and fruits	Tomatoes	KG	120.0	

1493 rows × 12 columns

In [12]:
```python
#Question3
dataframe.describe()
```

Out[12]:

	latitude	longitude	DzdPrice	usdprice	Year
count	1493.000000	1493.000000	1493.000000	1493.000000	1493.000000
mean	29.049174	-5.813241	167.447354	1.203295	2015.741460
std	3.569110	4.110959	151.740076	1.090434	1.094161
min	26.827448	-8.127653	4.000000	0.030000	2015.000000
25%	27.492310	-8.014000	60.000000	0.430000	2015.000000
50%	27.631001	-7.828405	110.000000	0.790000	2016.000000
75%	27.740000	-6.869948	230.000000	1.650000	2016.000000
max	36.752887	3.042048	700.000000	5.030000	2021.000000

In [13]:
```python
#Question4
#boxplot
_, bp = dataframe.boxplot( return_type='both')
```

Loading [MathJax]/extensions/Safe.js
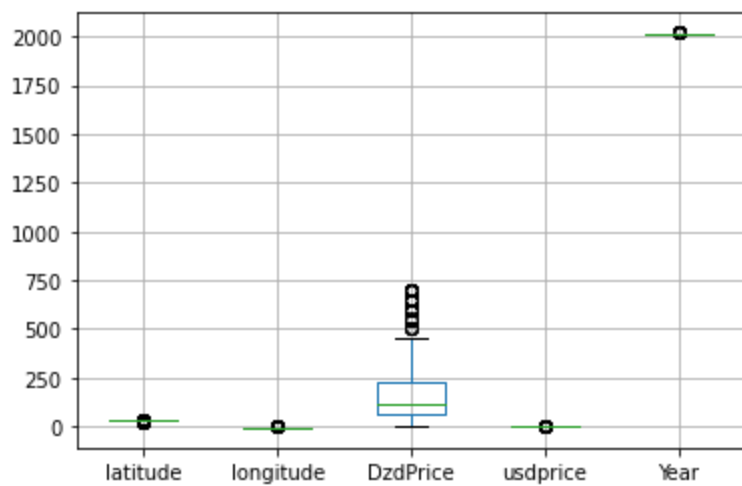
```
#Question5
dataframe.to_excel("output.xlsx")
```

```
#Data Visualisation

#Question1
print('Unique categories:\n',dataframe['category'].unique())
print('\n\nFrequent categories:\n',dataframe['category'].value_counts())
```

```
Unique categories:
 ['cereals and tubers' 'meat, fish and eggs' 'milk and dairy'
 'miscellaneous food' 'non-food' 'pulses and nuts' 'vegetables and fruits'
 'oil and fats']


Frequent categories:
 cereals and tubers       328
meat, fish and eggs      316
vegetables and fruits    212
pulses and nuts          166
miscellaneous food       154
milk and dairy           150
non-food                 145
oil and fats              22
Name: category, dtype: int64
```

```
#Question2
print('Unique cities:\n',dataframe['City'].unique())
print('\n\nFrequent of cities:\n',dataframe['City'].value_counts())
```

```
Unique cities:
 ['Hydra' 'Dakhla' 'Samara' 'Tindouf' 'Auserd' 'Layoun' 'Boujdour']


Frequent of cities:
 Tindouf     318
Samara      312
Dakhla      308
Hydra       262
Layoun      115
Auserd      106
Boujdour     72
Name: City, dtype: int64
```

```
#Question3
plt.rcParams["figure.figsize"] = (25, 5)
print("Movemonf of each commodity prices\n")
dataframe.plot(x='commodity', y='DzdPrice')
```

```
Movemonf of each commodity prices
```

Loading [MathJax]/extensions/Safe.js

`Out[18]:` `<AxesSubplot:xlabel='commodity'>`
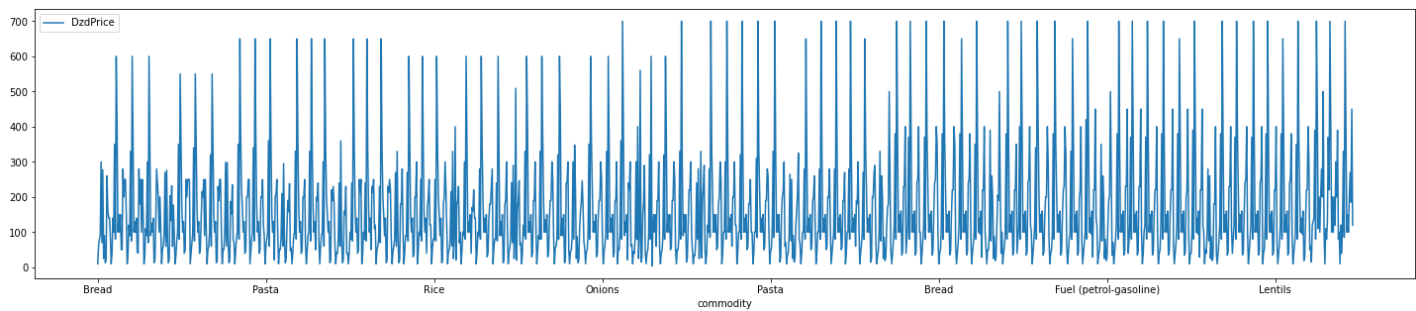

```
In [19]:  df_commodity = dataframe.set_index("commodity", drop = False)
          df_commodity
```

`Out[19]:`

commodity	date	State	City	latitude	longitude	category	commodity	unit	DzdPrice
Bread	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Bread	Unit	10.0
Pasta	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Pasta	500 G	60.0
Potatoes	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Potatoes	KG	80.0
Rice	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Rice	KG	90.0
Eggs	2015-04-15	Alger	Hydra	36.752887	3.042048	meat, fish and eggs	Eggs	30 pcs	300.0
...	...	...	...	...	...	...	...	...	...
Tea (sahm)	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	miscellaneous food	Tea (sahm)	Packet	200.0
Beans (white)	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	pulses and nuts	Beans (white)	KG	270.0
Lentils	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	pulses and nuts	Lentils	KG	185.0
Apples	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	vegetables and fruits	Apples	KG	450.0
Tomatoes	2021-05-15	Tindouf	Tindouf	27.676101	-8.127653	vegetables and fruits	Tomatoes	KG	120.0

1493 rows × 12 columns

```
In [20]:  print("Movement of Rice prices")
          df_commodity.loc["Rice",:].plot(x='commodity', y='DzdPrice')
```
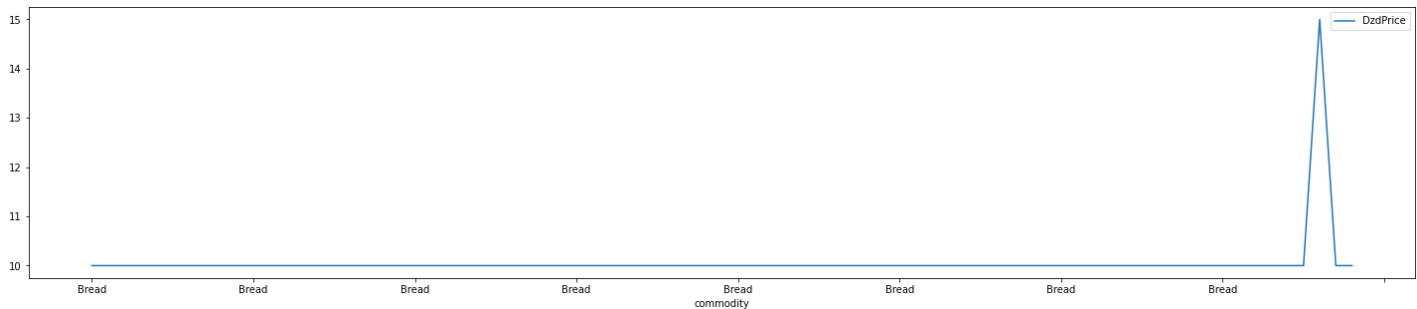
Movement of Rice prices

`Out[20]:` `<AxesSubplot:xlabel='commodity'>`

Loading [MathJax]/extensions/Safe.js
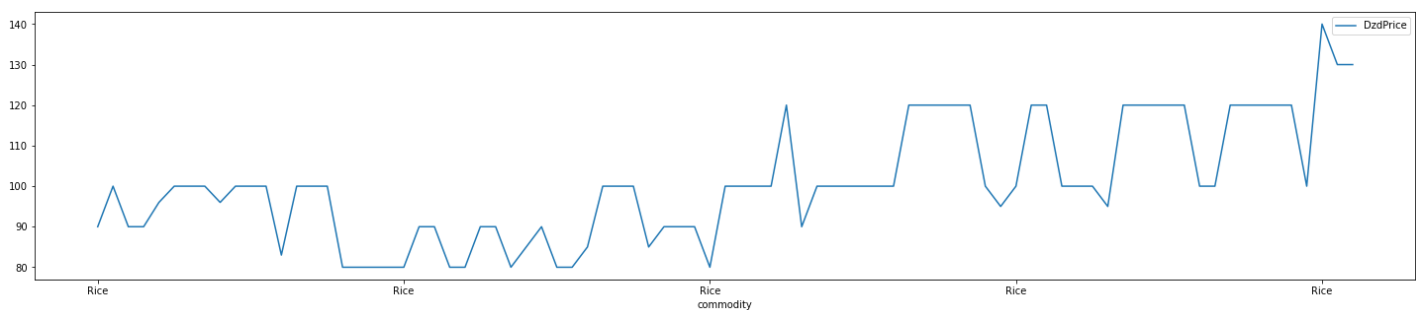
```
In [21]: print("Movement of Bread prices")
         df_commodity.loc["Bread",:].plot(x='commodity', y='DzdPrice')
```

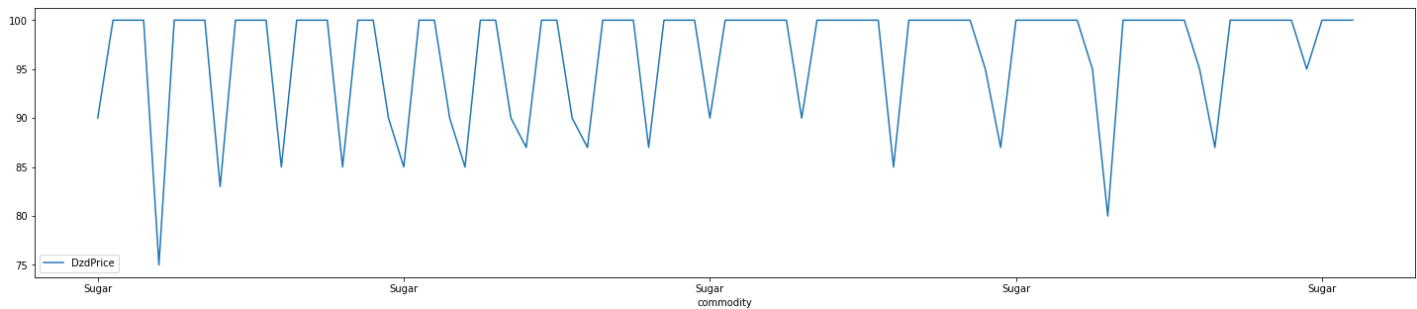Movement of Bread prices

Out[21]: <AxesSubplot:xlabel='commodity'>


```
In [22]: print("Movement of Sugar prices")
         df_commodity.loc["Sugar",:].plot(x='commodity', y='DzdPrice')
```

Movement of Sugar prices
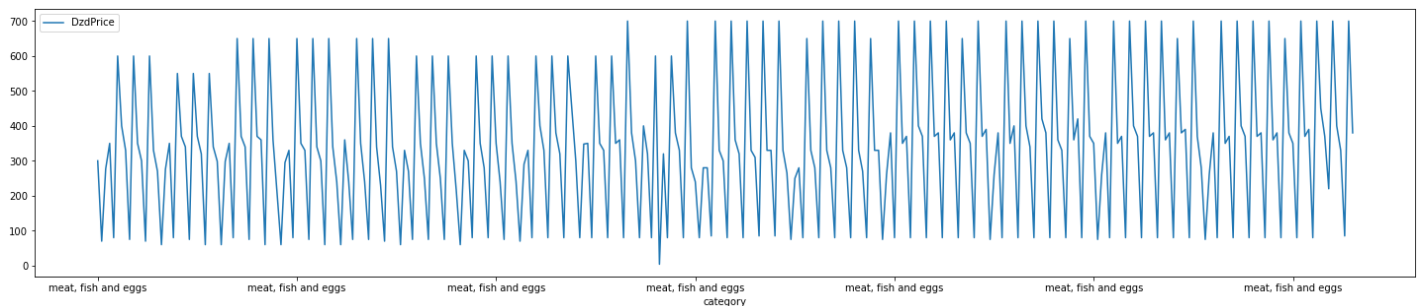
Out[22]: <AxesSubplot:xlabel='commodity'>


```
In [23]: #Question4
         df_category = dataframe.set_index("category", drop = False)
```

```
In [24]: print('Movemment of fish and eggs Category')
         df_category.loc["meat, fish and eggs",:].plot(x='category', y='DzdPrice')
```

Movemment of fish and eggs Category

Out[24]: <AxesSubplot:xlabel='category'>


```
In [25]: print('Movemment of cereals and tubers Category')
         df_category.loc["cereals and tubers",:].plot(x='category', y='DzdPrice')
```

Movemment of cereals and tubers Category

Out[25]: <AxesSubplot:xlabel='category'>

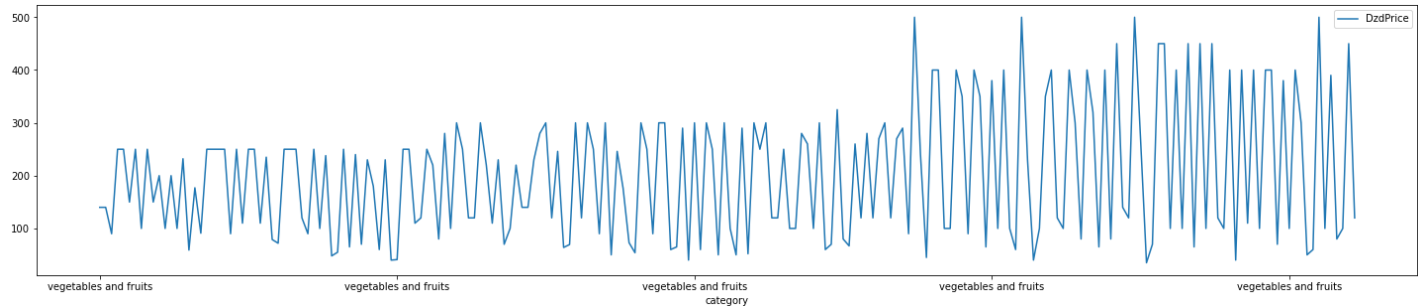Loading [MathJax]/extensions/Safe.js
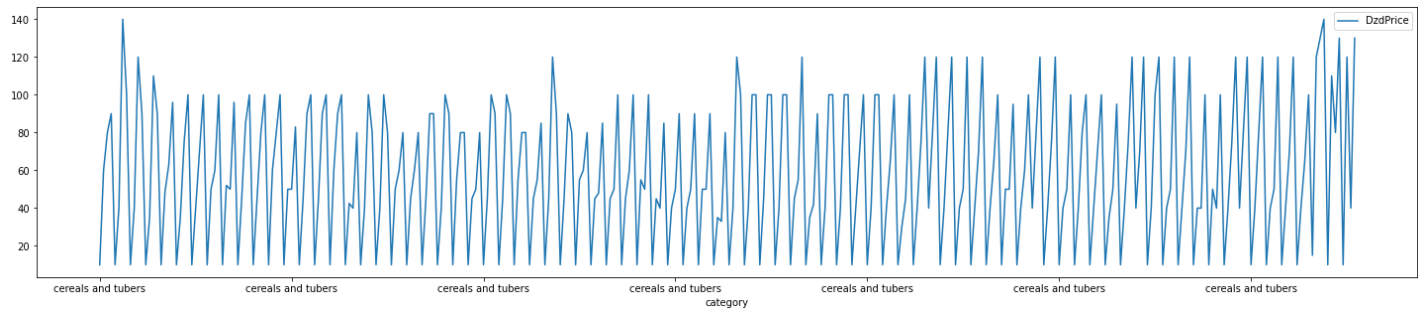
In [26]:

```
print('Movement of  vegetables and fruits Category')
df_category.loc["vegetables and fruits",:].plot(x='category', y='DzdPrice')
```
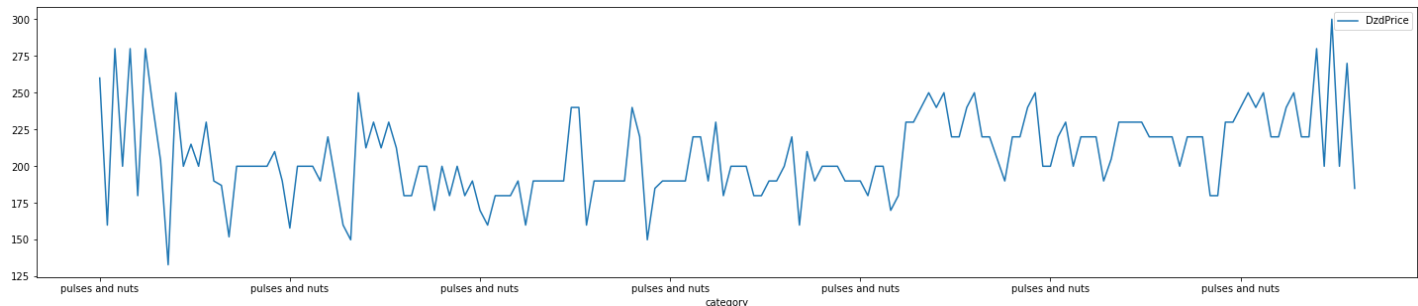
Movemment of vegetables and fruits Category

Out[26]: `<AxesSubplot:xlabel='category'>`

In [27]:

```
print('Movemment of pulses and nuts Category')
df_category.loc["pulses and nuts",:].plot(x='category', y='DzdPrice')
```

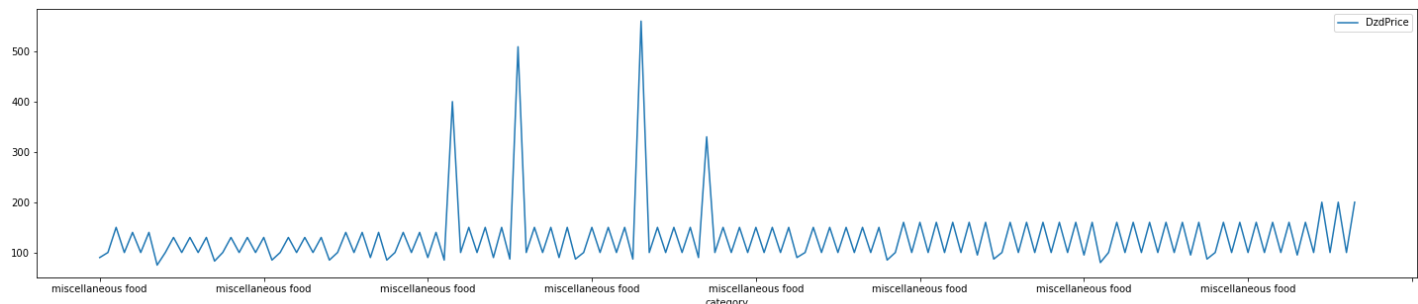Movemment of pulses and nuts Category

Out[27]: `<AxesSubplot:xlabel='category'>`

In [28]:

```
print('Movemment of miscellaneous food Category')
df_category.loc["miscellaneous food",:].plot(x='category', y='DzdPrice')
```
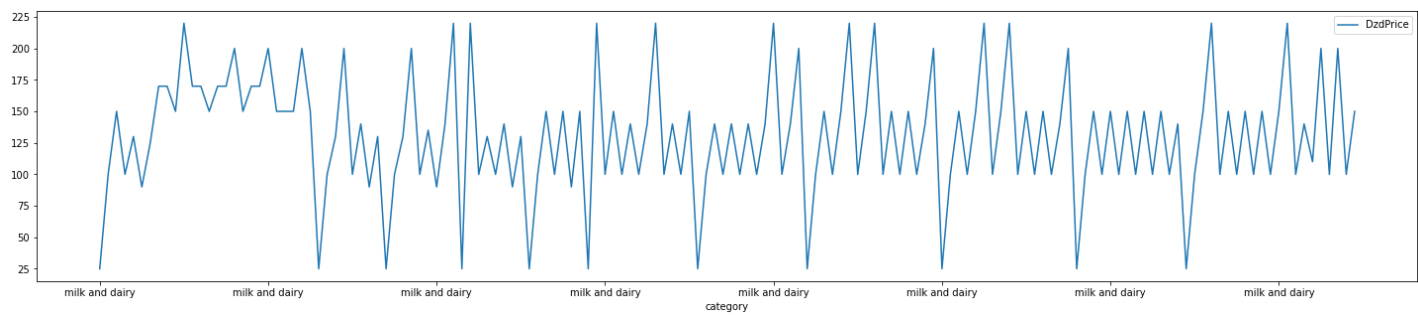
Movemment of miscellaneous food Category

Out[28]: `<AxesSubplot:xlabel='category'>`

In [29]:

```
print('Movemment of milk and dairy Category')
df_category.loc["milk and dairy",:].plot(x='category', y='DzdPrice')
```
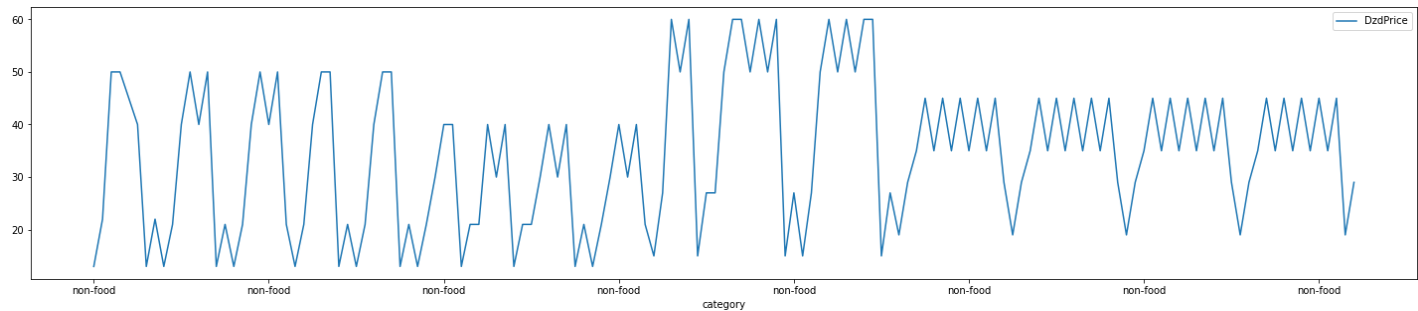
Movemment of milk and dairy Category

Loading [MathJax]/extensions/Safe.js

```
In [30]:  print('Movemment of non-food Category')
          df_category.loc["non-food",:].plot(x='category', y='DzdPrice')
```

Movemment of non-food Category
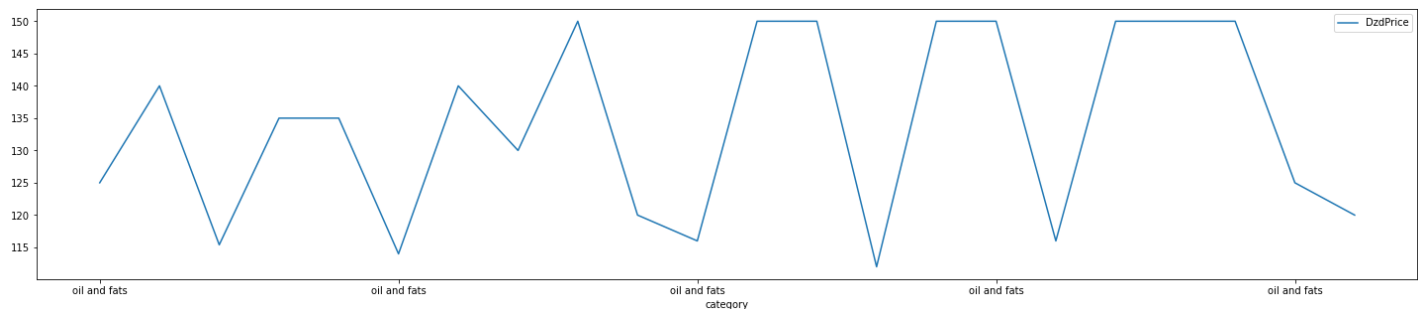
Out[30]: <AxesSubplot:xlabel='category'>


```
In [31]:  print('Movemment of oil anf fats Category')
          df_category.loc["oil and fats",:].plot(x='category', y='DzdPrice')
```

Movemment of oil anf fats Category

Out[31]: <AxesSubplot:xlabel='category'>


```
In [32]:  #Quetion5
          df_city = dataframe.set_index("City", drop = False)
          df_city_hydra = df_city.loc["Hydra", : ]
          df_city_hydra
```

Out[32]:

	date	State	City	latitude	longitude	category	commodity	unit	DzdPrice	usdprice
City										
Hydra	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Bread	Unit	10.0	0.07
Hydra	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Pasta	500 G	60.0	0.43
Hydra	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Potatoes	KG	80.0	0.57
Hydra	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Rice	KG	90.0	0.65
Hydra	2015-04-15	Alger	Hydra	36.752887	3.042048	meat, fish and eggs	Eggs	30 pcs	300.0	2.16

| City	date	State	City	latitude	longitude	category	commodity	unit	DzdPrice	usdprice	
...	...	...	...	...	...	...	...	...	...	...	...
Hydra	2016-10-15	Alger	Hydra	36.752887	3.042048	non-food	Fuel (petrol-gasoline)	L	29.0	0.21	2
Hydra	2016-10-15	Alger	Hydra	36.752887	3.042048	pulses and nuts	Beans (white)	KG	180.0	1.29	2
Hydra	2016-10-15	Alger	Hydra	36.752887	3.042048	pulses and nuts	Lentils	KG	180.0	1.29	2
Hydra	2016-10-15	Alger	Hydra	36.752887	3.042048	vegetables and fruits	Apples	KG	400.0	2.87	2
Hydra	2016-10-15	Alger	Hydra	36.752887	3.042048	vegetables and fruits	Tomatoes	KG	40.0	0.29	2

262 rows × 12 columns

```python
In [33]: df_city_hydra = df_city_hydra.filter(["commodity", "DzdPrice", "Month"])
         df_city_hydra = df_city_hydra.loc[df_city_hydra["commodity"] == "Eggs"]
         df_city_hydra
```

Out[33]:

City	commodity	DzdPrice	Month
Hydra	Eggs	300.0	4
Hydra	Eggs	270.0	5
Hydra	Eggs	298.5	6
Hydra	Eggs	210.0	7
Hydra	Eggs	240.0	8
Hydra	Eggs	270.0	9
Hydra	Eggs	215.0	10
Hydra	Eggs	241.0	11
Hydra	Eggs	300.0	12
Hydra	Eggs	300.0	1
Hydra	Eggs	240.0	2
Hydra	Eggs	265.0	3
Hydra	Eggs	330.0	7
Hydra	Eggs	390.0	8
Hydra	Eggs	350.0	9
Hydra	Eggs	280.0	10

```python
In [34]: df_city_hydra.plot.bar(x= 'Month', y='DzdPrice')
```

Out[34]: <AxesSubplot:xlabel='Month'>
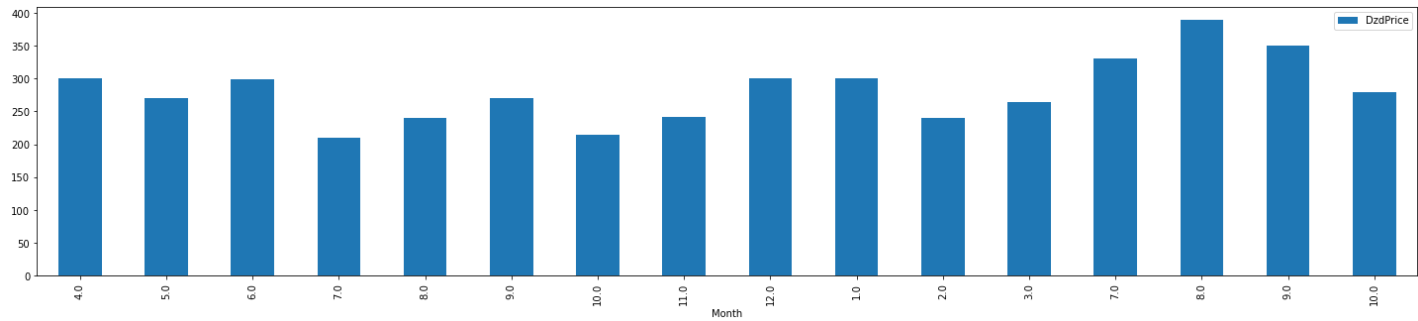
```python
#Question6
df_city_tindouf = df_city.loc["Tindouf", : ]
df_city_tindoufT = df_city_tindouf.filter(["commodity", "DzdPrice", "Month"])
df_city_tindoufT = df_city_tindoufT.loc[df_city_tindoufT["commodity"] == "Tea (sahm)"]
df_city_tindoufT
```

Out[35]:

City	commodity	DzdPrice	Month
Tindouf	Tea (sahm)	140.0	4
Tindouf	Tea (sahm)	130.0	5
Tindouf	Tea (sahm)	130.0	6
Tindouf	Tea (sahm)	130.0	7
Tindouf	Tea (sahm)	140.0	8
Tindouf	Tea (sahm)	140.0	9
Tindouf	Tea (sahm)	150.0	10
Tindouf	Tea (sahm)	150.0	11
Tindouf	Tea (sahm)	150.0	12
Tindouf	Tea (sahm)	150.0	1
Tindouf	Tea (sahm)	150.0	2
Tindouf	Tea (sahm)	150.0	3
Tindouf	Tea (sahm)	160.0	7
Tindouf	Tea (sahm)	160.0	8
Tindouf	Tea (sahm)	160.0	9
Tindouf	Tea (sahm)	160.0	10
Tindouf	Tea (sahm)	200.0	2021-05-15

In [36]:
```python
df_city_tindoufT.plot.bar(x= 'Month', y='DzdPrice')
```

Out[36]: <AxesSubplot:xlabel='Month'>

Loading [MathJax]/extensions/Safe.js
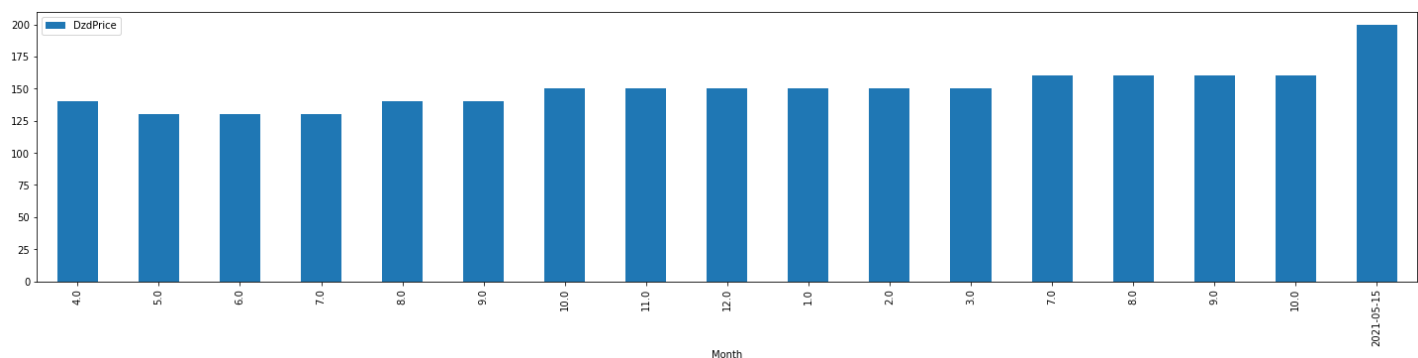
```
In [37]:   #Question7
           df_date = dataframe.set_index("date", drop = False)
           df_date= df_date["2015-04-15":"2015-12-15"]
           df_date
```

Out[37]:

date	date	State	City	latitude	longitude	category	commodity	unit	DzdPrice	usdprice
2015-04-15	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Bread	Unit	10.0	0.07
2015-04-15	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Pasta	500 G	60.0	0.43
2015-04-15	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Potatoes	KG	80.0	0.57
2015-04-15	2015-04-15	Alger	Hydra	36.752887	3.042048	cereals and tubers	Rice	KG	90.0	0.65
2015-04-15	2015-04-15	Alger	Hydra	36.752887	3.042048	meat, fish and eggs	Eggs	30 pcs	300.0	2.16
...	...	...	...	...	...	...	...	...	...	...
2015-12-15	2015-12-15	Tindouf	Tindouf	27.676101	-8.127653	pulses and nuts	Beans (white)	KG	240.0	1.72
2015-12-15	2015-12-15	Tindouf	Tindouf	27.676101	-8.127653	pulses and nuts	Lentils	KG	220.0	1.58
2015-12-15	2015-12-15	Tindouf	Tindouf	27.676101	-8.127653	vegetables and fruits	Apples	KG	300.0	2.16
2015-12-15	2015-12-15	Tindouf	Tindouf	27.676101	-8.127653	vegetables and fruits	Carrots	KG	60.0	0.43
2015-12-15	2015-12-15	Tindouf	Tindouf	27.676101	-8.127653	vegetables and fruits	Tomatoes	KG	65.0	0.47

636 rows × 12 columns

```
In [38]:   v = df_date['DzdPrice'].max()
           max_food = df_date.loc[df_date['DzdPrice'] == v]
           max_food['commodity']
```

Out[38]:   date
           2015-12-15   Meat (camel)
           Name: commodity, dtype: object

```
In [39]:   #Question8
           dataframe.plot(x = "DzdPrice" , y = "usdprice")
```

Out[39]:   <AxesSubplot:xlabel='DzdPrice'>