

```
In [6]: from pgmpy.models import BayesianModel
```

```
In [8]: cancer_model = BayesianModel([("Pollution", "Cancer"), ("Smoker", "Cancer"), ("Cancer", "Xray"), ("Cancer", "Dyspnoea")])
print(cancer_model)
```

```
In [9]: cancer_model.nodes()
```

```
Out[9]: NodeView(('Pollution', 'Cancer', 'Smoker', 'Xray', 'Dyspnoea'))
```

```
In [10]: cancer_model.edges()
```

```
Out[10]: OutEdgeView([('Pollution', 'Cancer'), ('Cancer', 'Xray'), ('Cancer', 'Dyspnoea'), ('Smoker', 'Cancer')])
```

```
In [11]: cancer_model.get_cpds()
```

```
Out[11]: []
```

```
In [12]: from pgmpy.factors.discrete import TabularCPD
```

```
In [13]: cpd_poll = TabularCPD(variable="Pollution", variable_card=2, values=[[0.9], [0.1]])
cpd_smoke = TabularCPD(variable="Smoker", variable_card=2, values=[[0.3], [0.7]])
cpd_cancer = TabularCPD(variable="Cancer", variable_card=2, values=[[0.03, 0.05, 0.07], [0.01, 0.02, 0.03]])
cpd_xray = TabularCPD(variable="Xray", variable_card=2, values=[[0.9, 0.2], [0.1, 0.8]])
cpd_dysp = TabularCPD(variable="Dyspnoea", variable_card=2, values=[[0.65, 0.3], [0.1, 0.9]])
cancer_model.add_cpds(cpd_poll, cpd_smoke, cpd_cancer, cpd_xray, cpd_dysp)
cancer_model.check_model()
```

```
Out[13]: True
```

```
In [14]: cancer_model.is_active_trail("Pollution", "Smoker")
```

```
Out[14]: False
```

```
In [15]: cancer_model.is_active_trail("Pollution", "Smoker", observed=["Cancer"])
```

```
Out[15]: True
```

In [16]: `cancer_model.get_cpds()`

Out[16]: [`<TabularCPD representing P(Pollution:2) at 0x2072d13ddc0>`,
`<TabularCPD representing P(Smoker:2) at 0x2072d13dd00>`,
`<TabularCPD representing P(Cancer:2 | Smoker:2, Pollution:2) at 0x20726d32820>`,
`<TabularCPD representing P(Xray:2 | Cancer:2) at 0x2072cfeb1f0>`,
`<TabularCPD representing P(Dyspnoea:2 | Cancer:2) at 0x20726d321f0>`]

In [17]: `print(cancer_model.get_cpds("Pollution"))`

```
+-----+-----+
| Pollution(0) | 0.9 |
+-----+-----+
| Pollution(1) | 0.1 |
+-----+-----+
```

In [18]: `print(cancer_model.get_cpds("Smoker"))`

```
+-----+-----+
| Smoker(0) | 0.3 |
+-----+-----+
| Smoker(1) | 0.7 |
+-----+-----+
```

In [19]: `print(cancer_model.get_cpds("Xray"))`

```
+-----+-----+-----+
| Cancer | Cancer(0) | Cancer(1) |
+-----+-----+-----+
| Xray(0) | 0.9       | 0.2       |
+-----+-----+-----+
| Xray(1) | 0.1       | 0.8       |
+-----+-----+-----+
```

In [20]: `print(cancer_model.get_cpds("Dyspnoea"))`

```
+-----+-----+-----+
| Cancer | Cancer(0) | Cancer(1) |
+-----+-----+-----+
| Dyspnoea(0) | 0.65      | 0.3      |
+-----+-----+-----+
| Dyspnoea(1) | 0.35      | 0.7      |
+-----+-----+-----+
```

In [21]: `print(cancer_model.get_cpds("Cancer"))`

Smoker	Smoker(0)	Smoker(0)	Smoker(1)	Smoker(1)
Pollution	Pollution(0)	Pollution(1)	Pollution(0)	Pollution(1)
Cancer(0)	0.03	0.05	0.001	0.02
Cancer(1)	0.97	0.95	0.999	0.98

In [22]: `cancer_model.local_independencies("Xray")`

Out[22]: (Xray _|_ Pollution, Smoker, Dyspnoea | Cancer)

In [23]: `cancer_model.local_independencies("Pollution")`

Out[23]: (Pollution _|_ Smoker)

In [24]: `cancer_model.local_independencies("Smoker")`

Out[24]: (Smoker _|_ Pollution)

In [25]: `cancer_model.local_independencies("Dyspnoea")`

Out[25]: (Dyspnoea _|_ Pollution, Smoker, Xray | Cancer)

In [26]: `cancer_model.local_independencies("Cancer")`

Out[26]:

```
In [27]: cancer_model.get_independencies()
```

```
Out[27]: (Pollution _|_ Smoker)
(Pollution _|_ Xray, Dyspnoea | Cancer)
(Pollution _|_ Xray, Dyspnoea | Smoker, Cancer)
(Pollution _|_ Xray | Dyspnoea, Cancer)
(Pollution _|_ Dyspnoea | Xray, Cancer)
(Pollution _|_ Xray | Smoker, Dyspnoea, Cancer)
(Pollution _|_ Dyspnoea | Smoker, Cancer, Xray)
(Smoker _|_ Pollution)
(Smoker _|_ Xray, Dyspnoea | Cancer)
(Smoker _|_ Xray, Dyspnoea | Pollution, Cancer)
(Smoker _|_ Dyspnoea | Xray, Cancer)
(Smoker _|_ Xray | Dyspnoea, Cancer)
(Smoker _|_ Dyspnoea | Pollution, Xray, Cancer)
(Smoker _|_ Xray | Pollution, Dyspnoea, Cancer)
(Xray _|_ Pollution, Smoker, Dyspnoea | Cancer)
(Xray _|_ Smoker, Dyspnoea | Pollution, Cancer)
(Xray _|_ Pollution, Dyspnoea | Smoker, Cancer)
(Xray _|_ Pollution, Smoker | Dyspnoea, Cancer)
(Xray _|_ Dyspnoea | Pollution, Smoker, Cancer)
(Xray _|_ Smoker | Pollution, Dyspnoea, Cancer)
(Xray _|_ Pollution | Smoker, Dyspnoea, Cancer)
(Dyspnoea _|_ Pollution, Smoker, Xray | Cancer)
(Dyspnoea _|_ Smoker, Xray | Pollution, Cancer)
(Dyspnoea _|_ Pollution, Xray | Smoker, Cancer)
(Dyspnoea _|_ Pollution, Smoker | Xray, Cancer)
(Dyspnoea _|_ Xray | Pollution, Smoker, Cancer)
(Dyspnoea _|_ Smoker | Pollution, Xray, Cancer)
(Dyspnoea _|_ Pollution | Smoker, Cancer, Xray)
```

```
In [28]: from pgmpy.inference import VariableElimination
```

```
In [29]: cancer_infr = VariableElimination(cancer_model)
```

```
In [30]: q = cancer_infr.query(variables=["Cancer"], evidence={"Smoker": 1})
print(q)
```

```
Finding Elimination Order: : 100%|██████████| 3/3 [00:00<00:00, 1503.87it/s]
Eliminating: Dyspnoea: 100%|██████████| 3/3 [00:00<00:00, 376.00it/s]
```

```
+-----+-----+
| Cancer | phi(Cancer) |
+-----+-----+
| Cancer(0) | 0.0029 |
+-----+-----+
| Cancer(1) | 0.9971 |
+-----+-----+
```

```
In [32]: q = cancer_infr.query(variables=["Cancer"], evidence={"Smoker": 1, "Pollution": 1})
print(q)
```

Finding Elimination Order: : 100%|██████████| 2/2 [00:00<00:00, 1003.78it/s]
 Eliminating: Dyspnoea: 100%|██████████| 2/2 [00:00<00:00, 501.62it/s]

```
+-----+-----+
| Cancer | phi(Cancer) |
+=====+=====+
| Cancer(0) | 0.0200 |
+-----+-----+
| Cancer(1) | 0.9800 |
+-----+-----+
```

extra code

```
In [33]: import numpy as np
from urllib.request import urlopen
import urllib
import matplotlib.pyplot as plt
import seaborn as sns
import sklearn as skl
import pandas as pd
```

```
In [35]: data= "http://archive.ics.uci.edu/ml/machine-learning-databases/heart-disease/proc
```

```
In [36]: names = ["age", "sex", "cp", "trestbps", "chol", "fbs", "restecg", "thalach", "ex
```

```
In [37]: heartDisease = pd.read_csv(urlopen(data), names=names)
heartDisease.head()
```

```
Out[37]:
```

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	heartdisei
0	28	1	2	130	132	0	2	185	0	0.0	?	?	?	
1	29	1	2	120	243	0	0	160	0	0.0	?	?	?	
2	29	1	2	140	?	0	0	170	0	0.0	?	?	?	
3	30	0	1	170	237	0	1	170	0	0.0	?	?	6	
4	31	0	2	100	219	0	1	150	0	0.0	?	?	?	

```
In [38]: del heartDisease["ca"]  
del heartDisease["slope"]  
del heartDisease["thal"]  
del heartDisease["oldpeak"]
```

```
In [39]: heartDisease = heartDisease.replace("?", np.nan)  
heartDisease.dtypes
```

```
Out[39]: age          int64  
sex          int64  
cp           int64  
trestbps     object  
chol         object  
fbs          object  
restecg      object  
thalach      object  
exang        object  
heartdisease int64  
dtype: object
```

```
In [40]: heartDisease.columns
```

```
Out[40]: Index(['age', 'sex', 'cp', 'trestbps', 'chol', 'fbs', 'restecg', 'thalach',  
               'exang', 'heartdisease'],  
              dtype='object')
```

```
In [41]: from pgmpy.models import BayesianModel  
from pgmpy.estimators import MaximumLikelihoodEstimator, BayesianEstimator
```

```
In [42]: model = BayesianModel([("age", "trestbps"), ("age", "fbs"), ("sex", "trestbps"), ("s
```

```
In [43]: model.fit(heartDisease, estimator=MaximumLikelihoodEstimator)
print(model.get_cpds("age"))
```

```
+-----+-----+
| age(28) | 0.00383142 |
+-----+-----+
| age(29) | 0.00383142 |
+-----+-----+
| age(30) | 0.00383142 |
+-----+-----+
| age(31) | 0.00766284 |
+-----+-----+
| age(32) | 0.0153257  |
+-----+-----+
| age(33) | 0.00766284 |
+-----+-----+
| age(34) | 0.0153257  |
+-----+-----+
| age(35) | 0.0191571  |
+-----+-----+
| age(36) | 0.0191571  |
+-----+-----+
| age(37) | 0.0306513  |
+-----+-----+
| age(38) | 0.0191571  |
+-----+-----+
| age(39) | 0.0344828  |
+-----+-----+
| age(40) | 0.0191571  |
+-----+-----+
| age(41) | 0.0383142  |
+-----+-----+
| age(42) | 0.0268199  |
+-----+-----+
| age(43) | 0.0421456  |
+-----+-----+
| age(44) | 0.0268199  |
+-----+-----+
| age(45) | 0.0229885  |
+-----+-----+
| age(46) | 0.045977   |
+-----+-----+
| age(47) | 0.0344828  |
+-----+-----+
| age(48) | 0.0613027  |
+-----+-----+
| age(49) | 0.0421456  |
+-----+-----+
| age(50) | 0.045977   |
+-----+-----+
| age(51) | 0.0344828  |
+-----+-----+
| age(52) | 0.0574713  |
+-----+-----+
| age(53) | 0.0383142  |
+-----+-----+
```

age(54)	0.0842912	
+-----+		+-----+
age(55)	0.0536398	
+-----+		+-----+
age(56)	0.0306513	
+-----+		+-----+
age(57)	0.0191571	
+-----+		+-----+
age(58)	0.0344828	
+-----+		+-----+
age(59)	0.0229885	
+-----+		+-----+
age(60)	0.00766284	
+-----+		+-----+
age(61)	0.00766284	
+-----+		+-----+
age(62)	0.00766284	
+-----+		+-----+
age(63)	0.00383142	
+-----+		+-----+
age(65)	0.0114943	
+-----+		+-----+
age(66)	0	
+-----+		+-----+

In [44]: `print(model.get_cpds("chol"))`

heartdisease	heartdisease(0)		heartdisease(1)	
+-----+		+-----+		+-----+
chol(100)	0.006134969325153374		0.0	
+-----+		+-----+		+-----+
chol(117)	0.0		0.01020408163265306	
+-----+		+-----+		+-----+
chol(129)	0.006134969325153374		0.0	
+-----+		+-----+		+-----+
chol(132)	0.006134969325153374		0.0	
+-----+		+-----+		+-----+
chol(147)	0.012269938650306749		0.0	
+-----+		+-----+		+-----+
chol(156)	0.0		0.01020408163265306	
+-----+		+-----+		+-----+
chol(160)	0.012269938650306749		0.01020408163265306	
+-----+		+-----+		+-----+
chol(161)	0.006134969325153374		0.0	
+-----+		+-----+		+-----+
chol(162)	0.006134969325153374		0.0	
+-----+		+-----+		+-----+

In [45]: `print(model.get_cpds("sex"))`

sex(0)	0.264368	
+-----+		+-----+
sex(1)	0.735632	
+-----+		+-----+


```
In [46]: model.get_independencies()
```

```
Out[46]: (age _|_ sex, exang)
(age _|_ exang | sex)
(age _|_ chol, thalach, restecg | heartdisease)
(age _|_ sex | exang)
(age _|_ sex, exang | fbs)
(age _|_ chol, thalach, restecg | heartdisease, sex)
(age _|_ exang | fbs, sex)
(age _|_ chol, thalach | heartdisease, restecg)
(age _|_ chol, thalach, restecg | heartdisease, exang)
(age _|_ chol, thalach, restecg | heartdisease, fbs)
(age _|_ chol, thalach, restecg | heartdisease, trestbps)
(age _|_ thalach, restecg | heartdisease, chol)
(age _|_ chol, restecg | heartdisease, thalach)
(age _|_ sex | fbs, exang)
(age _|_ heartdisease, thalach, restecg, chol | fbs, trestbps)
(age _|_ chol, thalach | heartdisease, sex, restecg)
(age _|_ chol, thalach, restecg | heartdisease, sex, exang)
(age _|_ chol, thalach, restecg | heartdisease, sex, fbs)
(age _|_ chol, thalach, restecg | heartdisease, trestbps, sex)
(age _|_ thalach, restecg | heartdisease, sex, chol)
```

```
In [47]: from pgmpy.inference import VariableElimination
heartDisease_infer = VariableElimination(model)
q = heartDisease_infer.query(variables=["heartdisease"],
evidence={"age": 29})

print(q)
```

```
Finding Elimination Order: : 100%|██████████| 7/7 [00:00<00:00, 3508.20it/s]
Eliminating: exang: 100%|██████████| 7/7 [00:00<00:00, 259.95it/s]
```

```
+-----+-----+
| heartdisease | phi(heartdisease) |
+=====+=====+
| heartdisease(0) | 0.6630 |
+-----+-----+
| heartdisease(1) | 0.3370 |
+-----+-----+
```

```
In [48]: q = heartDisease_infer.query(variables=["heartdisease"],
evidence={"chol": 100})

print(q)
```

C:\Users\User\anaconda3\lib\site-packages\pgmpy\factors\discrete\DiscreteFactor.py:438: UserWarning: Found unknown state name. Trying to switch to using all state names as state numbers

```
warn(
Finding Elimination Order: : 100%|██████████| 7/7 [00:00<00:00, 1400.64it/s]
Eliminating: exang: 100%|██████████| 7/7 [00:00<00:00, 143.24it/s]
```

```
+-----+-----+
| heartdisease | phi(heartdisease) |
+=====+=====+
| heartdisease(0) | 1.0000 |
+-----+-----+
| heartdisease(1) | 0.0000 |
+-----+-----+
```

In []: